



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 180133

TO: Ruixiang Li
Location: 4c75 / 4c70
Tuesday, February 28, 2006
Art Unit: 1646
Phone: 571-272-0875
Serial Number: 10 / 618570

From: Jan Delaval
Location: Biotech-Chem Library
Remsen 1a51
Phone: 571-272-2504

jan.delaval@uspto.gov

Search Notes

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

STIC-Biotech/ChemLib

180133

my

From: Li, Ruixiang
Sent: Tuesday, February 21, 2006 4:59 PM
To: STIC-Biotech/ChemLib
Subject: Sequence search of Application No.10/618570

Please do a standard search on:

SEQ ID NOS: 1 and 2 against nucleic acid databases (excluding pending databases).

Thank you very much!

Ruixiang Li
GAU 1646
REM 4D75
Mail Box 4C70
(571) 272-0875

STIC
2006
22

Searcher: Jan
Searcher Phone: 22504
Date Searcher Picked up: 2/22/06
Date completed: 2/23/06
Searcher Prep Time: 10
Online Time: 15

Type of Search
NA# ✓ AA#
S/L: Oligomer:
Encode/Transl: ✓
Structure #: Text:
Inventor: Litigation:

Vendors and cost where applicable
STN:
DIALOG:
QUESTEL/ORBIT:
LEXIS/NEXIS:
SEQUENCE SYSTEM: ✓
WWW/Internet:
Other (Specify):

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: February 16, 2006, 17:05:18 ; Search time 47 Seconds
(without alignments)
703.623 Million cell updates/sec

Title: US-10-618-570-2

Perfect score: 2059

Sequence: 1 MAQWDDPPQEDTDSCTES.....DKATRVGINFTLRTRQKE 400

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*

- 1: /cgn2_6/pctdata/1/aaa/5 COMB.pep.*
- 2: /cgn2_6/pctdata/1/aaa/5 COMB.pep.*
- 3: /cgn2_6/pctdata/1/aaa/5 COMB.pep.*
- 4: /cgn2_6/pctdata/1/aaa/5 COMB.pep.*
- 5: /cgn2_6/pctdata/1/aaa/5 COMB.pep.*
- 6: /cgn2_6/pctdata/1/aaa/5 COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|----------------------|
| 1 | 1381 | 67.1 | 349 | 6 | Patent No. 5510466 |
| 2 | 1381 | 67.1 | 453 | 6 | Patent No. 5510466-2 |
| 3 | 1047.5 | 50.9 | 451 | 1 | US-08-154-365-2 |
| 4 | 1042.5 | 50.6 | 358 | 1 | US-08-453-117-4 |
| 5 | 1042.5 | 50.6 | 358 | 1 | US-08-948-223-4 |
| 6 | 1042.5 | 50.6 | 358 | 1 | US-08-973-145-4 |
| 7 | 1042.5 | 50.6 | 358 | 1 | US-08-973-145-4 |
| 8 | 1042.5 | 50.6 | 358 | 1 | US-08-973-145-4 |
| 9 | 1042.5 | 50.6 | 358 | 1 | US-08-973-145-4 |
| 10 | 1042.5 | 50.6 | 358 | 1 | US-08-973-145-4 |
| 11 | 1042.5 | 50.6 | 358 | 1 | US-08-973-145-4 |
| 12 | 1042.5 | 50.6 | 358 | 1 | US-08-973-145-4 |
| 13 | 1042.5 | 50.6 | 358 | 1 | US-08-973-145-4 |
| 14 | 1042.5 | 50.6 | 358 | 1 | US-08-973-145-4 |
| 15 | 1042.5 | 50.6 | 358 | 1 | US-08-973-145-4 |
| 16 | 678 | 32.9 | 152 | 2 | US-08-831-393-4 |
| 17 | 678 | 32.9 | 152 | 2 | US-08-831-393-4 |
| 18 | 678 | 32.9 | 152 | 2 | US-08-831-393-4 |
| 19 | 138.5 | 6.7 | 159 | 2 | US-08-831-393-4 |
| 20 | 138.5 | 6.7 | 159 | 2 | US-08-831-393-4 |
| 21 | 138.5 | 6.7 | 159 | 2 | US-08-831-393-4 |
| 22 | 138.5 | 6.7 | 159 | 2 | US-08-831-393-4 |
| 23 | 138.5 | 6.7 | 159 | 2 | US-08-831-393-4 |
| 24 | 138.5 | 6.7 | 159 | 2 | US-08-831-393-4 |
| 25 | 138.5 | 6.7 | 159 | 2 | US-08-831-393-4 |
| 26 | 138.5 | 6.7 | 159 | 2 | US-08-831-393-4 |
| 27 | 138.5 | 6.7 | 159 | 2 | US-08-831-393-4 |

| | | | | | | |
|----|-------|-----|-----|---|-----------------|--------------------|
| 28 | 138.5 | 6.7 | 183 | 6 | 5168049-5 | Patent No. 5168049 |
| 29 | 138.5 | 6.7 | 415 | 1 | US-08-491-988-7 | Sequence 7, Appl |
| 30 | 138.5 | 6.7 | 421 | 1 | US-08-491-988-7 | Sequence 1, Appl |
| 31 | 138.5 | 6.7 | 421 | 1 | US-08-491-988-7 | Sequence 5, Appl |
| 32 | 138.5 | 6.6 | 128 | 1 | US-08-434-718-2 | Sequence 2, Appl |
| 33 | 138.5 | 6.6 | 128 | 1 | US-08-434-718-2 | Sequence 16, Appl |
| 34 | 138.5 | 6.6 | 128 | 2 | US-08-434-718-2 | Sequence 16, Appl |
| 35 | 138.5 | 6.6 | 128 | 2 | US-08-434-718-2 | Sequence 16, Appl |
| 36 | 138.5 | 6.6 | 128 | 2 | US-08-434-718-2 | Sequence 16, Appl |
| 37 | 138.5 | 6.6 | 128 | 2 | US-08-434-718-2 | Sequence 16, Appl |
| 38 | 138.5 | 6.6 | 128 | 2 | US-08-434-718-2 | Sequence 16, Appl |
| 39 | 138.5 | 6.6 | 128 | 2 | US-08-434-718-2 | Sequence 16, Appl |
| 40 | 138.5 | 6.6 | 128 | 2 | US-08-434-718-2 | Sequence 16, Appl |
| 41 | 138.5 | 6.6 | 128 | 2 | US-08-434-718-2 | Sequence 16, Appl |
| 42 | 138.5 | 6.6 | 128 | 2 | US-08-434-718-2 | Sequence 16, Appl |
| 43 | 138.5 | 6.6 | 128 | 2 | US-08-434-718-2 | Sequence 16, Appl |
| 44 | 138.5 | 6.6 | 128 | 2 | US-08-434-718-2 | Sequence 16, Appl |
| 45 | 138.5 | 6.6 | 128 | 2 | US-08-434-718-2 | Sequence 16, Appl |

ALIGNMENTS

RESULT 1
5510466-2
; Patent No. 5510466
; APPLICANT: KREIGER, MONTY; KODAMA, TATSUHIKO
; TITLE OF INVENTION: SCAVENGER RECEPTOR PROTEIN AND ANTIBODY
; THERE TO

NUMBER OF SEQUENCES: 12
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/307,400
FILING DATE: 16-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 997,113
FILING DATE: 24-DEC-1992
APPLICATION NUMBER: 391,486
FILING DATE: 09-AUG-1989
APPLICATION NUMBER: 272,002
FILING DATE: 15-NOV-1988
SEQ ID NO.2:
LENGTH: 349
5510466-2

| | | | | |
|-----------------------|----------------|--|----------|------------|
| Query Match | 67.1% | Score 1381 | DB 6 | Length 349 |
| Best Local Similarity | 100.0% | Pred. No. 4.3e-112 | | |
| Matches 272 | Conservative 0 | Mismatches 0 | Indels 0 | Gaps 0 |
| Qy | 1 | MAQWDDPPQEDTDSCTESVKEFDRSVTALLPPHPKNGPTLQERMSYKTAITLYLIV | 60 | |
| Db | 1 | MAQWDDPPQEDTDSCTESVKEFDRSVTALLPPHPKNGPTLQERMSYKTAITLYLIV | 60 | |
| Qy | 61 | FVVLVPIIGIVAAQLKWKTKCTVGSVNADISPSPEKNGSGEDMPFRAVMEKSMN | 120 | |
| Db | 61 | FVVLVPIIGIVAAQLKWKTKCTVGSVNADISPSPEKNGSGEDMPFRAVMEKSMN | 120 | |
| Qy | 121 | ESRQYLSNDEANLLDAKFNQFISITTDQRFNDVLPQLNSLSISQEHENIIGDISKSLV | 180 | |
| Db | 121 | ESRQYLSNDEANLLDAKFNQFISITTDQRFNDVLPQLNSLSISQEHENIIGDISKSLV | 180 | |
| Qy | 181 | GLNTTVLDLQPSITELNGRVQENAPKQSEMKLBERIYNASAEIKSLDKQVLEOEIK | 240 | |
| Db | 181 | GLNTTVLDLQPSITELNGRVQENAPKQSEMKLBERIYNASAEIKSLDKQVLEOEIK | 240 | |
| Qy | 241 | GEMKLLNNITNDLRKQWHSQTLKNITLLQ 272 | | |
| Db | 241 | GEMKLLNNITNDLRKQWHSQTLKNITLLQ 272 | | |

RESULT 2
5510466-4
; Patent No. 5510466
; APPLICANT: KREIGER, MONTY; KODAMA, TATSUHIKO

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 27, 2006, 08:21:52 ; Search time 18619 Seconds
(without alignments)
13009.116 Million cell updates/sec

Title: US-10-618-570-1
Perfect score: 5177
Sequence: 1 tttaagagaccaccaccgta.....agcgggggtcttcatttgg 5177

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_est3:*

4: gb_hic:*

5: gb_est4:*

6: gb_est5:*

7: gb_est6:*

8: gb_est7:*

9: gb_ges1:*

10: gb_ges2:*

11: gb_ges3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------|
| 1 | 829 | 16.0 | 935 | 7 | CK284786 |
| 2 | 812.4 | 15.7 | 856 | 7 | CK287297 |
| 3 | 812.4 | 15.7 | 910 | 7 | CK287930 |
| 4 | 812.4 | 15.7 | 933 | 7 | CK291799 |
| 5 | 812.4 | 15.7 | 936 | 7 | CK256977 |
| 6 | 812.4 | 15.7 | 947 | 7 | CK298208 |
| 7 | 812.4 | 15.7 | 954 | 7 | CK283361 |
| 8 | 803 | 15.5 | 804 | 7 | CK291519 |
| 9 | 794 | 15.3 | 811 | 7 | CK288185 |
| 10 | 793 | 15.3 | 793 | 8 | DN548862 |
| 11 | 737.4 | 14.2 | 767 | 8 | DR696294 |
| 12 | 707.2 | 13.7 | 949 | 6 | CF409555 |
| 13 | 634.2 | 12.3 | 837 | 10 | C2169992 |
| 14 | 605.4 | 11.7 | 878 | 7 | CK288911 |
| 15 | 589.4 | 11.4 | 591 | 8 | DR696887 |
| 16 | 588.8 | 11.4 | 789 | 9 | AQ361914 |
| 17 | 582.8 | 11.3 | 586 | 7 | CK801096 |
| 18 | 576.4 | 11.1 | 620 | 9 | AQ398387 |
| 19 | 566.6 | 10.9 | 1033 | 10 | AY413208 |
| 20 | 551.2 | 10.6 | 1027 | 10 | AY413209 |
| 21 | 542.4 | 10.5 | 549 | 7 | CK801630 |
| 22 | 541.2 | 10.5 | 804 | 8 | DR104979 |

| | | | | | |
|------|-------|------|-----|----|----------|
| 23 | 538.2 | 10.4 | 978 | 7 | CN643776 |
| 24 | 535 | 10.3 | 580 | 8 | DR108385 |
| 25 | 531.4 | 10.3 | 787 | 1 | AM037916 |
| c 26 | 530 | 10.2 | 759 | 6 | CB172376 |
| 27 | 523.8 | 10.1 | 637 | 9 | AQ47230 |
| 28 | 522.2 | 10.1 | 592 | 9 | AQ49162 |
| 29 | 514.8 | 9.9 | 561 | 9 | AQ447874 |
| 30 | 510.6 | 9.9 | 843 | 5 | EX496563 |
| c 31 | 508.4 | 9.8 | 562 | 6 | CB453499 |
| 32 | 508 | 9.8 | 509 | 9 | AQ447775 |
| 33 | 502.4 | 9.7 | 652 | 6 | CB453082 |
| 34 | 499 | 9.6 | 499 | 9 | AQ398880 |
| 35 | 496.4 | 9.6 | 498 | 8 | DR696295 |
| 36 | 494.2 | 9.6 | 498 | 9 | AQ398730 |
| c 37 | 484.2 | 9.5 | 857 | 9 | AQ875013 |
| 38 | 489.4 | 9.4 | 491 | 8 | DR696578 |
| 39 | 485.2 | 9.4 | 551 | 1 | AJ677598 |
| c 40 | 483.6 | 9.3 | 850 | 9 | AQ875052 |
| 41 | 483 | 9.3 | 707 | 8 | DR108377 |
| c 42 | 481.8 | 9.3 | 798 | 9 | AQ876139 |
| 43 | 479.6 | 9.3 | 740 | 10 | AG002307 |
| c 44 | 475.2 | 9.2 | 861 | 9 | AQ875040 |
| c 45 | 474.2 | 9.2 | 710 | 9 | BH972732 |

ALIGNMENTS

| | | | | |
|------------|--|--|--------|-----------------|
| RESULT 1 | CK284786 | 935 bp | linear | EST 02-AUG-2004 |
| LOCUS | EST747508 | Nicotiana benthamiana mixed tissue cDNA library, | | |
| DEFINITION | normalized, full-length Nicotiana benthamiana cDNA clone NBMAQ41 5' | | | |
| ACCESSION | CK284786 | | | |
| VERSION | CK284786.1 | GI:39858698 | | |
| KEYWORDS | EST. | | | |
| SOURCE | Nicotiana benthamiana | | | |
| ORGANISM | Nicotiana benthamiana | | | |
| REFERENCE | 1 (bases 1 to 935) | | | |
| AUTHORS | Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A., Day, B., Staskiewicz, B., Jin, H. and Baker, B. | | | |
| TITLE | Generation of EST sequences from Nicotiana benthamiana | | | |
| JOURNAL | Unpublished (2003) | | | |
| COMMENT | Other ESTs: EST747509 | | | |
| | Contact: Robin Buell | | | |
| | The Institute for Genomic Research | | | |
| | 9712 Medical Center Dr, Rockville, MD 20850, USA | | | |
| | Email: potato-array@tigr.org | | | |
| | Clones can be requested from the University of Arizona Genomics | | | |
| | Institute via http://genome.arizona.edu/orders/ | | | |
| | Seq primer: ATT TAG GTG ACA CTA TAG. | | | |
| FEATURES | Location/Qualifiers | | | |
| source | 1..935 | | | |
| | /organism="Nicotiana benthamiana" | | | |
| | /mol_type="mRNA" | | | |
| | /db_xref="taxon:4100" | | | |
| | /clones="NBMAQ41" | | | |
| | /tissue_type="abiotic and biotic stress-treated leaves, | | | |
| | callus tissue and root tissue" | | | |
| | /lab_host="DH10B-Tona" | | | |
| | /clone_lib="Nicotiana benthamiana mixed tissue cDNA | | | |
| | library, normalized, full-length" | | | |
| | /note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI; | | | |
| | supplier: RNA was isolated from Nicotiana benthamiana | | | |
| | tissues that include callus, roots from liquid culture | | | |
| | grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr), | | | |
| | cold-stressed leaves (5 C 3 hr, 6hr), and pathogen | | | |
| | challenged leaves (Pseudomonas syringae pv tomato 12 hr; | | | |
| | Xanthomonas campestris pv campestris 12 hr, 18hr; | | | |

Pseudomonas syringae pv *phaseolicola* 18hr, and *Xanthomonas campestris* pv *vesicatoria* 18hr). RNA was isolated from these tissues and pooled in approximately equal molar amounts."

| | |
|---------------------------|---|
| ORIGIN | |
| Query Match | 16.0%; Score 829; DB 7; Length 935; |
| Best Local Similarity | 99.4%; Pred. No. 1.5e-215; |
| Matches 832; Conservative | 0; Mismatches 5; Indels 0; Gaps 0; |
| QY | 3706 ACAGATGAGATCGTTTCGATGATGAACAAGATGATGCACGCGAGGTTCTCCGGCC 3765 |
| Db | 11 ACCGGATCTGGATCGTTTCGATGATGAACAAGATGATGCACGCGAGGTTCTCCGGCC 70 |
| QY | 3766 GCTTGGGTGGAGAGCTATTTCGGCTATCATCTGGGCAACAACAGACATCGGCTGCTGAT 3825 |
| Db | 71 GCTTGGGTGGAGAGCTATTTCGGCTATCATCTGGGCAACAACAGACATCGGCTGCTGAT 130 |
| QY | 3826 GCGCGCGTTCGGCTGTCAGCGCAGGGCGCGCGTTCCTTTTGTCAAGACCGACCTG 3885 |
| Db | 131 GCGCGCGTTCGGCTGTCAGCGCAGGGCGCGCGTTCCTTTTGTCAAGACCGACCTG 190 |
| QY | 3886 TCCGCTGCCCTGAATGAATCTCAGACGAGGCGCGGCTATCGTGGCTGGCCACGACG 3945 |
| Db | 191 TCCGCTGCCCTGAATGAATCTCAGACGAGGCGCGGCTATCGTGGCTGGCCACGACG 250 |
| QY | 3946 GCGTTCCTTGGCAGCTGTGCTGACGTTGTCACTGAAGCGGAGAGGACTGCTGCTA 4005 |
| Db | 251 GCGTTCCTTGGCAGCTGTGCTGACGTTGTCACTGAAGCGGAGAGGACTGCTGCTA 310 |
| QY | 4006 TTGGSCGAAGTGGCGGGCAGGATCTCTGTCTCATCTCACTTGTCTCTGCCGGAAGTA 4065 |
| Db | 311 TTGGSCGAAGTGGCGGGCAGGATCTCTGTCTCATCTCACTTGTCTCTGCCGGAAGTA 370 |
| QY | 4066 TCCATCATGCTGATGCAATCGCGGCTGTGATGATCGGCTACCTGCGCCATTC 4125 |
| Db | 371 TCCATCATGCTGATGCAATCGCGGCTGTGATGATCGGCTACCTGCGCCATTC 430 |
| QY | 4126 GACCAACGAGCAACATCGCATCGAGCGAGCAGTACTCGGATGGAAGCGGCTTGTGTC 4185 |
| Db | 431 GACCAACGAGCAACATCGCATCGAGCGAGCAGTACTCGGATGGAAGCGGCTTGTGTC 490 |
| QY | 4186 GATCAGGATGATCTGGACGAAAGACATCAGGGGCTCGCGCCAGCGCACTGTCGCCAGG 4245 |
| Db | 491 GATCAGGATGATCTGGACGAAAGACATCAGGGGCTCGCGCCAGCGCACTGTCGCCAGG 550 |
| QY | 4246 CTCAGGCGCGCATGCCGACGGGATCTCGTGTGATCCCATGGCGATGCTGCTGTTG 610 |
| Db | 551 CTCAGGCGCGCATGCCGACGGGATCTCGTGTGATCCCATGGCGATGCTGCTGTTG 610 |
| QY | 4306 CGAATATCATGTTGGAATAATGGCGGCTTTTCTGATTCATCGACTGTGGCGGCTGGGT 4365 |
| Db | 611 CGAATATCATGTTGGAATAATGGCGGCTTTTCTGATTCATCGACTGTGGCGGCTGGGT 670 |
| QY | 4366 GTGGCGGACCGTATCAGACATACGCTTGGCTACCCGTGATATGCTGAAGAGCTTGGC 4425 |
| Db | 671 GTGGCGGACCGTATCAGACATACGCTTGGCTACCCGTGATATGCTGAAGAGCTTGGC 730 |
| QY | 4426 GCGCAATGGGCTGACCGCTTCTCGTGTCTTACGGTATCGCGGCTCCCGATTCGACGCG 4485 |
| Db | 731 GCGCAATGGGCTGACCGCTTCTCGTGTCTTACGGTATCGCGGCTCCCGATTCGACGCG 790 |
| QY | 4486 ATCGGCTTCTATCGGCTTCTGAAGATTCCTTCTGAGCGGAGCTCTGGGTTTCGATA 4542 |
| Db | 791 ATCGGCTTCTATCGGCTTCTTACGAGTCTTCTTACGCGGAGCTCTGGGTTTCGATA 847 |

RESULT 2
CK287297 856 bp mRNA linear EST 02-AUG-2004
LOCUS
DEFINITION
EST750019 *Nicotiana benthamiana* mixed tissue cDNA library,
normalized, full-length *Nicotiana benthamiana* cDNA clone NBM815 5'
end, mRNA sequence.

CK287297
CK287297.1 GI:39863696
EST.
SOURCE
ORGANISM
Nicotiana benthamiana
Nicotiana benthamiana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; asterids; lamiales; Solanales; Solanaceae; *Nicotiana*.
1 (bases 1 to 856)
Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A., Day, B., Staskawicz, B., Jin, H. and Baker, B.
Generation of EST sequences from *Nicotiana benthamiana* Unpublished (2003)
Other ESTs: EST750020
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
Clones can be requested from the University of Arizona Genomics Institute via <http://genome.arizona.edu/orders/>.
Seq primer: ATT TAG GTG ACA CTA TAG.

FEATURES
source
1. 856
/organism="Nicotiana benthamiana"
/mol_type="mRNA"
/db_xref="taxon:4100"
/clone="NBM815"
/tissue_type="abiotic and biotic stress-treated leaves, callus tissue and root tissue"
/lab_host="DH10B-Tona"
/clone_lib="Nicotiana benthamiana mixed tissue cDNA library, normalized, full-length"
/notes="Vector: pCMVSPORT6.1; Site_1: EcoRI; Site_2: NotI; supplier: RNA was isolated from *Nicotiana benthamiana* tissues that include callus, roots from liquid culture grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr), cold-stressed leaves (5 C 3 hr, 6hr), and pathogen challenged leaves (*Pseudomonas syringae* pv tomato 12 hr; *Xanthomonas campestris* pv *campestris* 12 hr, 18hr; *Xanthomonas syringae* pv *phaseolicola* 18hr, and *Xanthomonas campestris* pv *vesicatoria* 18hr). RNA was isolated from these tissues and pooled in approximately equal molar amounts."

| | |
|---------------------------|---|
| ORIGIN | |
| Query Match | 15.7%; Score 812.4; DB 7; Length 856; |
| Best Local Similarity | 99.9%; Pred. No. 5.2e-211; |
| Matches 813; Conservative | 0; Mismatches 1; Indels 0; Gaps 0; |
| QY | 3729 GATTGAACAAGATGGATTGCACGAGGTTCTCCGGCCGCTTGGGTGAGAGGCTATTTCGG 3788 |
| Db | 14 GATTGAACAAGATGGATTGCACGAGGTTCTCCGGCCGCTTGGGTGAGAGGCTATTTCGG 73 |
| QY | 3789 CTATGACTGGGCACAACAGACAATCGGCTGCTCTGATGCCGCCGTGTTCGGGCTGTGAGC 3848 |
| Db | 74 CTATGACTGGGCACAACAGACAATCGGCTGCTCTGATGCCGCCGTGTTCGGGCTGTGAGC 133 |
| QY | 3849 GCAGGGCGCGCGGTTCTTTTGTCAAGACGCGACCTGTCGGTGCCCTGATGAATGCA 3908 |
| Db | 134 GCAGGGCGCGCGGTTCTTTTGTCAAGACGCGACCTGTCGGTGCCCTGATGAATGCA 193 |
| QY | 3909 GGACGAGCGAGCGGCTATCGTGGTGGCCACGACGGGGGTTCTTCGCGAGCTGTGCT 3968 |
| Db | 194 GGACGAGCGAGCGGCTATCGTGGTGGCCACGACGGGGGTTCTTCGCGAGCTGTGCT 253 |
| QY | 3969 CGAGCTGTGTCACTGAAGCGGGAAGGACTGGCTGCTATTGGGCGAAGTTCGCGGGCGAGGA 4028 |
| Db | 254 CGAGCTGTGTCACTGAAGCGGGAAGGACTGGCTGCTATTGGGCGAAGTTCGCGGGCGAGGA 313 |
| QY | 4029 TCTCCTCTCATCTCCTTCTCTCTGCGGAAAGTATCCATCATGCTGATGCAATGCG 4089 |
| Db | 314 TCTCCTCTCATCTCCTTCTCTCTGCGGAAAGTATCCATCATGCTGATGCAATGCG 373 |

QY 4089 CGGGCTGATACGCTTGATCCGGCTACCTGCCATTGCAACCAAGCGAATCATCGCAT 4148
Db |||||||
QY 374 CGGGCTGATACGCTTGATCCGGCTACCTGCCATTGCAACCAAGCGAATCATCGCAT 433
Db |||||||
QY 4149 CGAGCGACGATGCTACTCGGATGGAAGCGGCTTCTGTGATCAGATGATCTGGACGAAGA 4208
Db |||||||
QY 434 CGAGCGACGATGCTACTCGGATGGAAGCGGCTTCTGTGATCAGATGATCTGGACGAAGA 493
Db |||||||
QY 4209 GCATCAGGGGCTCGCGCCAGCGCACTGTTCCGCAAGGCTCAAGCGCGCATGCGCGACGG 4268
Db |||||||
QY 494 GCATCAGGGGCTCGCGCCAGCGCACTGTTCCGCAAGGCTCAAGCGCGCATGCGCGACGG 553
Db |||||||
QY 4269 CGAGGATCTGCTGATGACCCATGCGGATGCTGCTGCGGATCATGCTGGAATGG 4328
Db |||||||
QY 554 CGAGGATCTGCTGATGACCCATGCGGATGCTGCTGCGGATCATGCTGGAATGG 613
Db |||||||
QY 4329 CCGCTTTTCTGGATTCATGCACTGTGGCGCGGCTGGGTGCGGACCGCTATCAGGACAT 4388
Db |||||||
QY 614 CCGCTTTTCTGGATTCATGCACTGTGGCGCGGCTGGGTGCGGACCGCTATCAGGACAT 673
Db |||||||
QY 4389 AGCGTTGGCTACCGGTGATATGCTGAAGAGCTTGGCGGGAATGGGCTGACCGCTTCT 4448
Db |||||||
QY 674 AGCGTTGGCTACCGGTGATATGCTGAAGAGCTTGGCGGGAATGGGCTGACCGCTTCT 733
Db |||||||
QY 4449 CGTGTTCATCGGTATCGCGCTCCGATTCGAGGCGATCGCTTCTATCGCTTCTTGA 4508
Db |||||||
QY 734 CGTGTTCATCGGTATCGCGCTCCGATTCGAGGCGATCGCTTCTATCGCTTCTTGA 793
Db |||||||
QY 4509 CGAGTTCTTCGAGCGGACTCTGGGTTTCGATA 4542
Db |||||||
QY 794 CGAGTTCTTCGAGCGGACTCTGGGTTTCGATA 827
Db |||||||

RESULT 3

CK287930
LOCUS EST750652 Nicotiana benthamiana mixed tissue cDNA library,
normalized, full-length Nicotiana benthamiana cDNA clone NBMBCT5 5',
end, mRNA sequence.

ACCESSION CK287930.1 GI:39864940

VERSION EST.

KEYWORDS

SOURCE

ORGANISM

Nicotiana benthamiana

Nicotiana benthamiana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;

asterids; lamiales; Solanales; Solanaceae; Nicotiana.

1 (bases 1 to 910)

Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A., Day, B.,

Staskiewicz, B., Jin, H., and Baker, B.

Generation of EST sequences from Nicotiana benthamiana

Unpublished (2003)

Other ESTs: EST750653

Contact: Robin Buell

The Institute for Genomic Research

9712 Medical Center Dr, Rockville, MD 20850, USA

Email: potato-array@tigr.org

Clones can be requested from the University of Arizona Genomics

Institute via http://genome.arizona.edu/orders/

Seq primer: Art TAG GTG ACA CTA TAG.

Location/Qualifiers

1. 910

/organism="Nicotiana benthamiana"

/mol_type="mRNA"

/db_xref="taxon:4100"

/clone="NBMBCT5"

/tissue_type="abiotic and biotic stress-treated leaves,"

callus tissue and root tissue"

/lab_host="DH10B-TonA"

/clone_lib="Nicotiana benthamiana mixed tissue cDNA

library, normalized, full-length"

/note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;

supplier: RNA was isolated from Nicotiana benthamiana

source

FEATURES

source

RESULT 4

RESULT 4

RESULT 4

RESULT 4

RESULT 4

RESULT 4

CK291799
 LOCUS EST754513 Nicotiana benthamiana mixed tissue cDNA library, linear EST 02-AUG-2004
 DEFINITION normalized, full-length Nicotiana benthamiana cDNA clone NEMC477 5' end, mRNA sequence.

ACCESSION CK291799
 VERSION CK291799.1 GI:39872608
 KEYWORDS EST.
 SOURCE Nicotiana benthamiana
 ORGANISM Nicotiana benthamiana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; asterids; lamids; Solanales; Solanaceae; Nicotiana.

REFERENCE 1 (bases 1 to 933)
 AUTHORS Buell C.R., Hart, A., Zismann, V., Karamycheva, S.A., Day, B., Staskawicz, B., Jin, H. and Baker, B.
 TITLE Generation of EST sequences from Nicotiana benthamiana
 JOURNAL Unpublished (2003)
 COMMENT Other ESTs: EST754514
 Contact: Robin Buell
 The Institute for Genomic Research
 9712 Medical Center Dr, Rockville, MD 20850, USA
 Email: potato-array@tigr.org
 Clones can be requested from the University of Arizona Genomics Institute via <http://genome.arizona.edu/orders/>.
 Seq primer: ATT TAG GTG ACA CTA TAG.

FEATURES
 Location/Qualifiers
 1..933
 /organism="Nicotiana benthamiana"
 /mol_type="mRNA"
 /db_xref="taxon:4100"
 /clone="NEMC477"
 /tissue_type="abiotic and biotic stress-treated leaves, callus tissue and root tissue"
 /lab_hosts="DH10B-Tona"
 /clone_lib="Nicotiana benthamiana mixed tissue cDNA library, normalized, full-length"
 /note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI; supplier: RNA was isolated from Nicotiana benthamiana tissues that include callus, roots from liquid culture grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr), cold-stressed leaves (5 C 3 hr, 6hr), and pathogen challenged leaves (Pseudomonas syringae pv tomato 12 hr; Xanthomonas campestris pv campestris 12 hr, 18hr; Pseudomonas syringae pv phaseolicola 18hr, and Xanthomonas campestris pv vesicatoria 18hr). RNA was isolated from these tissues and pooled in approximately equal molar amounts."

ORIGIN
 Query Match 15.7%; Score 812.4; DB 7; Length 933;
 Best Local Similarity 99.9%; Pred. No. 5.3e-211;
 Matches 813; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3729 GATTGAACAAGATGGATTGCACGAGGTTCCTCCGGCGCTTGGGTGGAGAGGCTATTTCGG 3788
 DB |||||||
 QY 3789 CTATGACTGGGCAACAACAGCAATCGGCTGCTCTGATGCCCGCTGTTCGGCTGTCCAGC 3848
 DB |||||||
 QY 106 CTATGACTGGGCAACAACAGCAATCGGCTGCTCTGATGCCCGCTGTTCGGCTGTCCAGC 165
 DB |||||||
 QY 3849 GCAGGGGGCCCGGCTCTTTTGTTCAGACCGACTGTCGGTCCCTGATGACTGCA 3908
 DB |||||||
 QY 166 GCAGGGGGCCCGGCTCTTTTGTTCAGACCGACTGTCGGTCCCTGATGACTGCA 225
 DB |||||||
 QY 3909 GCAGGAGCGACGGGCTATCGTGGTGGCCACGACGGCGCTTCCTTGGCAGCTGTGCT 3968
 DB |||||||
 QY 226 GGAAGGACGACGGGCTATCGTGGTGGCCACGACGGCGCTTCCTTGGCAGCTGTGCT 285
 DB |||||||
 QY 3969 CGAGCTGTCTACTGAAGCGGAAGGACTGCTGCTATTGGGCGAAGTGGCGGGCAGGA 4028
 DB |||||||
 QY 286 CGAGCTGTCTACTGAAGCGGAAGGACTGCTGCTATTGGGCGAAGTGGCGGGCAGGA 345
 DB |||||||

QY 4029 TCTCCTGTCTACCTTCTCTCTGCGGAGAAAGTATCCATCATGCTGATGCAATGCG 4088
 DB |||||||
 QY 346 TCTCCTGTCTACCTTCTCTCTGCGGAGAAAGTATCCATCATGCTGATGCAATGCG 405
 DB |||||||
 QY 4089 GCGGCTGCATACGCTTGTGATCCGGCTACTCTGCCCATTCGACCAACGAAACATCGCAT 4148
 DB |||||||
 QY 406 GCGGCTGCATACGCTTGTGATCCGGCTACTCTGCCCATTCGACCAACGAAACATCGCAT 465
 DB |||||||
 QY 4149 CGAGCGAGCAGTACTCTCGGATGGAGCGGCTCTTGTGATCAGGATCATCTGGACGAGA 4208
 DB |||||||
 QY 466 CGAGCGAGCAGTACTCTCGGATGGAGCGGCTCTTGTGATCAGGATCATCTGGACGAGA 525
 DB |||||||
 QY 4209 GCATCAGGGGCTCGCGCAGCGAACTGTTGCGCAGGCTCAAGCGCGCATGCCCGACGG 4268
 DB |||||||
 QY 526 GCATCAGGGGCTCGCGCAGCGAACTGTTGCGCAGGCTCAAGCGCGCATGCCCGACGG 585
 DB |||||||
 QY 4269 CGAGGATCTCGTCTGATGCCCATCGCGATGCTGCTGCCGAATATCATGTGGAAATGG 4328
 DB |||||||
 QY 586 CGAGGATCTCGTCTGATGCCCATCGCGATGCTGCTGCCGAATATCATGTGGAAATGG 645
 DB |||||||
 QY 4329 CCGCTTTTCTGGATTCTCATCGACTGTGCGCGGCTGGGTGGCGGACCGCTATCAGGACAT 4388
 DB |||||||
 QY 646 CCGCTTTTCTGGATTCTCATCGACTGTGCGCGGCTGGGTGGCGGACCGCTATCAGGACAT 705
 DB |||||||
 QY 4389 AGCGTTGGCTACCGCTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCCT 4448
 DB |||||||
 QY 706 AGCGTTGGCTACCGCTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCCT 765
 DB |||||||
 QY 4449 CGTGTCTTACGGTATCGCGCTCCCGATTCGACGCGCATCGCCTTCTATCGCCTTCTTGA 4508
 DB |||||||
 QY 766 CGTGTCTTACGGTATCGCGCTCCCGATTCGACGCGCATCGCCTTCTATCGCCTTCTTGA 825
 DB |||||||
 QY 4509 CGAGTTCTTCTGACGGGACTCTGGGGTTCGATA 4542
 DB |||||||

RESULT 5
 CK256977
 LOCUS EST740614 potato callus cDNA library, linear EST 30-JUL-2004
 DEFINITION Solanum tuberosum cDNA clone POCD170 5' end, mRNA sequence.

ACCESSION CK256977
 VERSION CK256977.1 GI:39813957
 KEYWORDS EST.
 SOURCE Solanum tuberosum (potato)
 ORGANISM Solanum tuberosum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; asterids; lamids; Solanales; Solanaceae; Solanum.

REFERENCE 1 (bases 1 to 936)
 AUTHORS Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A. and Baker, B.
 TITLE Generation of ESTs from potato callus tissue
 JOURNAL Unpublished (2003)
 COMMENT Contact: Robin Buell
 The Institute for Genomic Research
 9712 Medical Center Dr, Rockville, MD 20850, USA
 Email: potato-array@tigr.org
 Clones can be requested from the University of Arizona Genomics Institute via <http://genome.arizona.edu/orders/>.
 Seq primer: ATT TAG GTG ACA CTA TAG.

FEATURES
 Location/Qualifiers
 1..936
 /organism="Solanum tuberosum"
 /mol_type="mRNA"
 /cultivar="Kennebec"
 /db_xref="taxon:4113"
 /clone="POCD170"
 /tissue_type="callus"
 /lab_hosts="DH10B-Tona"
 /clone_lib="potato callus cDNA library, normalized and full-length"

/note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;
supplier: RNA was isolated from Solanum tuberosum var.
Kennebec callus tissue grown on solid media."

```
ORIGIN
Query Match      15.7%; Score 812.4; DB 7; Length 936;
Best Local Similarity 99.9%; Pred. No. 5.3e-211;
Matches 813; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3729 GATTGAACAAGATGGATTGACGAGGTTCTCCGCGCGCTTGGGTGGAGAGGCTATTTCGG 3788
Db 14 GATTGAACAAGATGGATTGACGAGGTTCTCCGCGCGCTTGGGTGGAGAGGCTATTTCGG 73

Qy 3789 CTATGACTGGGCAACAACAGCAATCGGCTGCTCTGATGCGCGCGCTGTTCCGGCTGTTCAGC 3848
Db 74 CTATGACTGGGCAACAACAGCAATCGGCTGCTCTGATGCGCGCGCTGTTCCGGCTGTTCAGC 133

Qy 3849 GCAGGGGCGCGCGTCTCTTTTGTCAACACCGACTGTCCGGTGCCTGTAATGAACATGCA 3908
Db 134 GCAGGGGCGCGCGTCTCTTTTGTCAACACCGACTGTCCGGTGCCTGTAATGAACATGCA 193

Qy 3909 GGAGGAGCAGCGCGGCTATCGTGGCTGGCCACGACGGCGCTTCTTTCGCGAGCTGTGCT 3968
Db 194 GGAGGAGCAGCGCGGCTATCGTGGCTGGCCACGACGGCGCTTCTTTCGCGAGCTGTGCT 253

Qy 3969 CGACGTTGTCTAAGCGGGAAGGACTGGCTGCTATTTGGGCGAAGTGGCGGGGCGGAGGA 4028
Db 254 CGACGTTGTCTAAGCGGGAAGGACTGGCTGCTATTTGGGCGAAGTGGCGGGGCGGAGGA 313

Qy 4029 TCTCTGTCTATCTACCTTGTCTTCTGCGGAGAAAGTATCCATGCTGCTGATGCAATGCG 4088
Db 314 TCTCTGTCTATCTACCTTGTCTTCTGCGGAGAAAGTATCCATGCTGCTGATGCAATGCG 373

Qy 4089 GCGGCTGCATACGCTTGTATCGCGGTACTCTGCGCCATTCGACCACCAAGCGAAATCGCAT 4148
Db 374 GCGGCTGCATACGCTTGTATCGCGGTACTCTGCGCCATTCGACCACCAAGCGAAATCGCAT 433

Qy 4149 CGAGCGAGCAGTACTCGATGGAAGCGGCTTGTTCGATCAGATGATCTGGACGAAGA 4208
Db 434 CGAGCGAGCAGTACTCGATGGAAGCGGCTTGTTCGATCAGATGATCTGGACGAAGA 493

Qy 4209 GCATCAGGGGCTCGGCGGCGGAACTTTCGCCAGGCTCAAGCGGCGCATGCCGCGAGG 4268
Db 494 GCATCAGGGGCTCGGCGGCGGAACTTTCGCCAGGCTCAAGCGGCGCATGCCGCGAGG 553

Qy 4269 CGAGGATCTCTGTCGACCCATGCGATGCGCTGCTTTCGCGAATATCATGTGGAAATGG 4328
Db 554 CGAGGATCTCTGTCGACCCATGCGATGCGCTGCTTTCGCGAATATCATGTGGAAATGG 613

Qy 4329 CCGCTTTTCTGGATTTCATCGACTGTGGCGGCTGGGTGTGGCGGACCGCTATCAGGACAT 4388
Db 614 CCGCTTTTCTGGATTTCATCGACTGTGGCGGCTGGGTGTGGCGGACCGCTATCAGGACAT 673

Qy 4389 AGCTTGGCTACCGTGATATGCTGAAGAGCTTGGCGGCGAATATGAGGCTGACCGCTTCT 4448
Db 674 AGCTTGGCTACCGTGATATGCTGAAGAGCTTGGCGGCGAATATGAGGCTGACCGCTTCT 733

Qy 4449 CGTCTTTTACGGTATCGCGCTCCGATTCGACGCGCATCGCTTCTATCGCCTTCTTGA 4508
Db 734 CGTCTTTTACGGTATCGCGCTCCGATTCGACGCGCATCGCTTCTATCGCCTTCTTGA 793

Qy 4509 CGAGTTCTTCTGAGCGGAGCTCTGGGGTTCGATA 4542
Db 794 CGAGTTCTTCTGAGCGGAGCTCTGGGGTTCGATA 827
```

RESULT 6
CK298208
LOCUS
DEFINITION EST760922 Nicotiana benthamiana mixed tissue cDNA library,
normalized, full-length Nicotiana benthamiana cDNA clone NEMD30 5'
end, mRNA sequence.
CK298208
ACCESSION

VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

CK298208.1 GI:39885354
EST.
Nicotiana benthamiana
Nicotiana benthamiana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamids; Solanales; Solanaceae; Nicotiana.
1 (bases 1 to 947)
Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A., Day, B.,
Staskawicz, B., Jin, H. and Baker, B.
Generation of EST sequences from Nicotiana benthamiana
Unpublished (2003)
Other ESTs: EST760923
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
Clones can be requested from the University of Arizona Genomics
Institute via http://genome.arizona.edu/orders/
Seq primer: ATT TAG GTG ACA CTA TAG.

FEATURES
source

1. .947
Location/Qualifiers
/organism="Nicotiana benthamiana"
/mol_type="mRNA"
/db_xref="taxon:4100"
/clone="NEMD30"
/tissue_type="abiotic and biotic stress-treated leaves,
callus tissue and root tissue"
/lab_host="DH10B-Tona"
/clone_lib="Nicotiana benthamiana mixed tissue cDNA
library, normalized, full-length"
/notes="vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;
supplier: RNA was isolated from Nicotiana benthamiana
tissues that include callus, roots from liquid culture
grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr),
cold-stressed leaves (5 C 3 hr, 6hr), and pathogen
challenged leaves (Pseudomonas syringae pv tomato 12 hr;
Xanthomonas campestris pv campestris 12 hr, 18hr;
Pseudomonas syringae pv phaseolicola 18hr, and Xanthomonas
campestris pv vesicatoria 18hr). RNA was isolated from
these tissues and pooled in approximately equal molar
amounts."

ORIGIN

Query Match 15.7%; Score 812.4; DB 7; Length 947;
Best Local Similarity 99.9%; Pred. No. 5.3e-211;
Matches 813; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 3729 GATTGAACAAGATGGATTGACGAGGTTCTCCGCGCGCTTGGGTGGAGAGGCTATTTCGG 3788
Db 17 GATTGAACAAGATGGATTGACGAGGTTCTCCGCGCGCTTGGGTGGAGAGGCTATTTCGG 76
Qy 3789 CTATGACTGGGCAACAACAGCAATCGGCTGCTCTGATGCGCGCGCTGTTCCGGCTGTTCAGC 3848
Db 77 CTATGACTGGGCAACAACAGCAATCGGCTGCTCTGATGCGCGCGCTGTTCCGGCTGTTCAGC 136
Qy 3849 GCAGGGGCGCGCGTCTCTTTTGTCAAGACCGACTGTCCGGTGCCTGTAATGAACATGCA 3908
Db 137 GCAGGGGCGCGCGTCTCTTTTGTCAAGACCGACTGTCCGGTGCCTGTAATGAACATGCA 196
Qy 3909 GGAGGAGCAGCGCGGCTATCGTGGCTGGCCACGACGGCGCTTCTTTCGCGAGCTGTGCT 3968
Db 197 GGAGGAGCAGCGCGGCTATCGTGGCTGGCCACGACGGCGCTTCTTTCGCGAGCTGTGCT 256
Qy 3969 CGAGTTGTCTAAGCGGGAAGGAGCTGGCTGCTATTTGGGCGAAGTGGCGGGGCGGAGGA 4028
Db 257 CGAGTTGTCTAAGCGGGAAGGAGCTGGCTGCTATTTGGGCGAAGTGGCGGGGCGGAGGA 316
Qy 4029 TCTCTGTCTATCTACCTTGTCTGCGGAGAAAGTATCCATGCTGCTGATGCAATGCG 4088
Db 317 TCTCTGTCTATCTACCTTGTCTGCGGAGAAAGTATCCATGCTGCTGATGCAATGCG 376
Qy 4089 GCGGCTGCATACGCTTGTGATCCGGCTACCTGCCCATTCGACCACCAAGCGAAATCGCAT 4148

|||||
Db 377 CGCGCTGATACGCTTGTACCGCTACCTGCCATTCGACCAACGAAACATCGCAT 436
|||||
Qy 4149 CGAGGAGCAGTACTCGGATGGAAGCCGGTCTTGTGATCAGATGATCTGGAGGAAGA 4208
|||||
Db 437 CGAGGAGCAGTACTCGGATGGAAGCCGGTCTTGTGATCAGATGATCTGGAGGAAGA 496
|||||
Qy 4209 GCATCAGGGGCTCGCGCAGCCGAACTCTTCGACAGGCTCAAGGCGCGCATGCCGACGG 4268
|||||
Db 497 GCATCAGGGGCTCGCGCAGCCGAACTCTTCGACAGGCTCAAGGCGCGCATGCCGACGG 556
|||||
Qy 4269 CGAGGATCTCGTGGTACCCATGCGGATGCTGCTTGCCTGATATCATGTGGGAAATGG 4328
|||||
Db 557 CGAGGATCTCGTGGTACCCATGCGGATGCTGCTTGCCTGATATCATGTGGGAAATGG 616
|||||
Qy 4329 CGCTTTTCTGGATTCATGATGCGGCTGGCGGCTGGGTGGCGGACCGCTATCAGGACAT 4388
|||||
Db 617 CGCTTTTCTGGATTCATGATGCGGCTGGCGGCTGGGTGGCGGACCGCTATCAGGACAT 676
|||||
Qy 4389 AGCGTTGGCTACCGTGTATATGCTGAAGAGCTTGGCGGCAATGGGCTGACCGCTTCCT 4448
|||||
Db 677 AGCGTTGGCTACCGTGTATATGCTGAAGAGCTTGGCGGCAATGGGCTGACCGCTTCCT 736
|||||
Qy 4449 CGTGTCTTACGATATCGCGCTCCCGATTCGAGCGCATCGCTTCTATCGCTTCTTGA 4508
|||||
Db 737 CGTGTCTTACGATATCGCGCTCCCGATTCGAGCGCATCGCTTCTATCGCTTCTTGA 796
|||||
Qy 4509 CGAGTTCTTCGAGCGGACATCTGGGGTTCGATA 4542
|||||
Db 797 CGAGTTCTTCGAGCGGACATCTGGGGTTCGAAA 830
|||||

RESULT 7

CK283361

LOCUS EST746083 Nicotiana benthamiana mixed tissue cDNA library,
DEFINITION normalized, full-length Nicotiana benthamiana cDNA clone NBWAG50 5'
end, mRNA sequence.

ACCESSION

CK283361.1 GI:3985898

VERSION

EST.

KEYWORDS

Nicotiana benthamiana

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamiales; Solanales; Solanaceae; Nicotiana.

REFERENCE

1 (bases 1 to 954)

Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A., Day, B.,

Staskiewicz, B., Jin, H. and Baker, B.

Generation of EST sequences from Nicotiana benthamiana

Unpublished (2003)

Other ESTs: EST746084

Contact: Robin Buell

The Institute for Genomic Research

9712 Medical Center Dr. Rockville, MD 20850, USA

Email: potato-array@tigr.org

Clones can be requested from the University of Arizona Genomics

Institute via <http://genome.arizona.edu/orders/>.

Seq primer: ATT TAG GTG ACA CTA TAG.

Location/Qualifiers

1. 954

/organism="Nicotiana benthamiana"

/mol_type="mRNA"

/db_xref="taxon:4100"

/clone="NBWAG50"

/tissue_type="abiotic and biotic stress-treated leaves,

callus tissue and root tissue"

/lab_host="DH10B-TonA"

/clone_lib="Nicotiana benthamiana mixed tissue cDNA

library, normalized, full-length"

/note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;

supplier: RNA was isolated from Nicotiana benthamiana

tissues that include callus, roots from liquid culture

grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr),
cold-stressed leaves (5 C 3 hr, 6hr), and pathogen
challenged leaves (Pseudomonas syringae pv tomato 12 hr;
Xanthomonas campestris pv phaseolicola 18hr, 18hr;
Pseudomonas syringae pv phaseolicola 18hr, and Xanthomonas
campestris pv vesicatoria 18hr). RNA was isolated from
these tissues and pooled in approximately equal molar
amounts."

ORIGIN

Query Match 15.7%; Score 812.4; DB 7; Length 954;
Best Local Similarity 99.9%; Pred. No. 5.3e-211;
Matches 813; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 3729 GATTGAACAGATGGATTGACGCAAGGTTCTCCGGCCCGCTTGGGTGAGAGGCTATTTCGG 3788
Db 14 GATTGAACAGATGGATTGACGCAAGGTTCTCCGGCCCGCTTGGGTGAGAGGCTATTTCGG 73
Qy 3789 CTATGACTGGGCACAAACAGACAAATCGGCTGCTCTGATGCGCGCGCTGTTCGGCTGTGAGC 3848
Db 74 CTATGACTGGGCACAAACAGACAAATCGGCTGCTCTGATGCGCGCGCTGTTCGGCTGTGAGC 133
Qy 3849 GCAGGGCGCCCGGCTCTTTTGTCAAGACCGACCTCTCCGGTGCCTCTGAATGAATGCA 3908
Db 134 GCAGGGCGCCCGGCTCTTTTGTCAAGACCGACCTCTCCGGTGCCTCTGAATGAATGCA 193
Qy 3909 GGAGGAGGACCGCGGCTATCGTGGCTGGCCACGACGGGCGTTCCTTGGCAGCTGTGCT 3968
Db 194 GGAGGAGGACCGCGGCTATCGTGGCTGGCCACGACGGGCGTTCCTTGGCAGCTGTGCT 253
Qy 3969 CGACGTTGTCACTGAAGCGGAAGGAGTGGCTGCTATTGGGCGAAGTGGCGGGCAGGA 4028
Db 254 CGACGTTGTCACTGAAGCGGAAGGAGTGGCTGCTATTGGGCGAAGTGGCGGGCAGGA 313
Qy 4029 TCTCTGTCTATCTCACCTTCTCTCGCGAAGATATCCATCATATGCTGATCAATGCG 4088
Db 314 TCTCTGTCTATCTCACCTTCTCTCGCGAAGATATCCATCATATGCTGATCAATGCG 373
Qy 4089 CGCGCTGCATACGCTTGCATCGGCTACCTGCCCATTTGCACCAACCAAGCAACATCGCAT 4148
Db 374 CGCGCTGCATACGCTTGCATCGGCTACCTGCCCATTTGCACCAACCAAGCAACATCGCAT 433
Qy 4149 CGAGCGAGCAGCTACTCGGATGGAAGCGGCTCTTGTGATCAGATGATCTGGACGAAGA 4208
Db 434 CGAGCGAGCAGCTACTCGGATGGAAGCGGCTCTTGTGATCAGATGATCTGGACGAAGA 493
Qy 4209 GCATCAGGGGCTCGCGCAGCCGAACTGTTTCGCCAGGCTCAAGGCGCGCATGCCGACGG 4268
Db 494 GCATCAGGGGCTCGCGCAGCCGAACTGTTTCGCCAGGCTCAAGGCGCGCATGCCGACGG 553
Qy 4269 CGAGGATCTCGTGTACCCATGGCGATGCTGCTGCCGAATATCATGTGGGAAATGG 4328
Db 554 CGAGGATCTCGTGTACCCATGGCGATGCTGCTGCCGAATATCATGTGGGAAATGG 613
Qy 4329 CCGCTTTTCTGGATTATCATCGACTGTGGCCGGCTGGGTGGCGACCGCTATCAGGACAT 4388
Db 614 CCGCTTTTCTGGATTATCATCGACTGTGGCCGGCTGGGTGGCGACCGCTATCAGGACAT 673
Qy 4389 AGCGTTGGCTACCCGCTGATATTCCTGAAGAGCTTGGCGGCGAAATGGGCTGACCGCTTCCT 4448
Db 674 AGCGTTGGCTACCCGCTGATATTCCTGAAGAGCTTGGCGGCGAAATGGGCTGACCGCTTCCT 733
Qy 4449 CGTGTCTTACGGTATGCGCGCTCCCGATTCGACGCGCATCGCTTCTATCGCTTCTTGA 4508
Db 734 CGTGTCTTACGGTATGCGCGCTCCCGATTCGACGCGCATCGCTTCTATCGCTTCTTGA 793
Qy 4509 CGAGTTCTTCTGAGCGGGACTCTGGGGTTCGATA 4542
Db 794 CGAGTTCTTCTGAGCGGGACTCTGGGGTTCGAAA 827
|||||

RESULT 8

CK291519

LOCUS CK291519 804 bp mRNA linear EST 02-AUG-2004
 DEFINITION EST754233 Nicotiana benthamiana mixed tissue cDNA library,
 normalized, full-length Nicotiana benthamiana cDNA clone NEMC276 5',
 end, mRNA sequence.
 ACCESSION CK291519
 VERSION CK291519.1 GI:39872047
 KEYWORDS EST.
 SOURCE Nicotiana benthamiana
 ORGANISM Nicotiana benthamiana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 asterids; lamids; Solanales; Solanaceae; Nicotiana.
 REFERENCE 1 (bases 1 to 804)
 AUTHORS Buell,C.R., Hart,A., Zismann,V., Karamycheva,S.A., Day,B.,
 Staskawicz,B., Jin,H. and Baker,B.
 TITLE Generation of EST sequences from Nicotiana benthamiana
 JOURNAL Unpublished (2003)
 COMMENT Contact: Robin Buell
 The Institute for Genomic Research
 9712 Medical Center Dr, Rockville, MD 20850, USA
 Email: potato-array@tigr.org
 Clones can be requested from the University of Arizona Genomics
 Institute via <http://genome.arizona.edu/orders/>
 Seq primer: ATT TAG GTG ACA CTA TAG.
 FEATURES
 Location/Qualifiers
 1..804
 /organism="Nicotiana benthamiana"
 /mol_type="mRNA"
 /db_xref="taxon:4100"
 /clone="NEMC276"
 /tissue_type="abiotic and biotic stress-treated leaves,
 callus tissue and root tissue"
 /lab_host="DH10B-Tona"
 /clone_lib="Nicotiana benthamiana mixed tissue cDNA
 library, normalized, full-length"
 /note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;
 supplier: RNA was isolated from Nicotiana benthamiana
 tissues that include callus, roots from liquid culture
 grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr),
 cold-stressed leaves (5 C 3 hr, 6hr), and pathogen
 challenged leaves (Pseudomonas syringae pv tomato 12 hr;
 Xanthomonas campestris pv campestris 12 hr, 18hr;
 Pseudomonas syringae pv phaseolicola 18hr, and Xanthomonas
 campestris pv vesicatoria 18hr). RNA was isolated from
 these tissues and pooled in approximately equal molar
 amounts."
 ORIGIN
 Query Match 15.5%; Score 803; DB 7; Length 804;
 Best Local Similarity 100.0%; Pred. No. 1.9e-208;
 Matches 803; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3729 GATTGAACAGATGATTCACGAGGTTCTCCGCGCTGGGTGGAGAGGCTATTCCG 3788
 DB 2 GATTGAACAGATGATTCACGAGGTTCTCCGCGCTGGGTGGAGAGGCTATTCCG 61
 QY 3789 CTATGACTGGGCAACACAGACATCGGCTGCTGATGCCCGCTGTTCCGGCTGTCCAGC 3848
 DB 62 CTATGACTGGGCAACACAGACATCGGCTGCTGATGCCCGCTGTTCCGGCTGTCCAGC 121
 QY 3849 CGAGGGGCGCGGTTCTTTTGTCAAGACCGACTGTCGGTGCCTGGAATGAATGCA 3908
 DB 122 CGAGGGGCGCGGTTCTTTTGTCAAGACCGACTGTCGGTGCCTGGAATGAATGCA 181
 QY 3909 GGAGGAGCGAGCGGCTATCGTGGCTGGCCACGACGGCGCTTCTTGGCAGCTGTGCT 3968
 DB 182 GGAGGAGCGAGCGGCTATCGTGGCTGGCCACGACGGCGCTTCTTGGCAGCTGTGCT 241
 QY 3969 CGAGCTTCTCACTGAAGCGGAAGGACTGCTGCTATTGGCGCAAGTGGCGGCGCAGGA 4028
 DB 242 CGAGCTTCTCACTGAAGCGGAAGGACTGCTGCTATTGGCGCAAGTGGCGGCGCAGGA 301
 QY 4029 TCTCCTGTGTCATCTCACCTTGTCTCTGCGGAGAAAGTATCCATCATGCTGATGCAATGCG 4088

Db 302 TCTCCTGTGTCATCTCACCTTGTCTTCTGCGGAGAAAGTATCCATCATGGCTGATGCAATCGG 361
 QY 4089 GCGGCTGCATACGCTTGTATCGGCTACTCTGCCATTTGACACCAACGAAGCAATCGCAT 4148
 Db 362 GCGGCTGCATACGCTTGTATCGGCTACTCTGCCATTCGACCAACCAAGCAATCGCAT 421
 QY 4149 CGAGCGAGCAGTACTCGATGGAAGCGGCTGTTGTCGATCAGATGATCTGGACGAAGA 4208
 Db 422 CGAGCGAGCAGTACTCGATGGAAGCGGCTGTTGTCGATCAGATGATCTGGACGAAGA 481
 QY 4209 GCATCAGGGGCTCGCGCAGCGCAACTGTTCCGAGGCTCAAGCGCGCATGCCGACGG 4268
 Db 482 GCATCAGGGGCTCGCGCAGCGCAACTGTTCCGAGGCTCAAGCGCGCATGCCGACGG 541
 QY 4269 CGAGGATCTCGTCTGACCCATGCGGCTGCTGTCGGAATATCATGTTGGAAAAATGG 4328
 Db 542 CGAGGATCTCGTCTGACCCATGCGGCTGCTGTCGGAATATCATGTTGGAAAAATGG 601
 QY 4329 CCGCTTTTCTGGATTCATCGACTGTGGCGGCTGGGTGGCGGACCGCTATCAGGACAT 4388
 Db 602 CCGCTTTTCTGGATTCATCGACTGTGGCGGCTGGGTGGCGGACCGCTATCAGGACAT 661
 QY 4389 AGCGTTGCTACCGCTGATATGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCT 4448
 Db 662 AGCGTTGCTACCGCTGATATGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCT 721
 QY 4449 CGTCTTTTACGGTATCGCGCTCCCGATTCGACGCGATTCGCTTCTATCGCTTCTTGA 4508
 Db 722 CGTCTTTTACGGTATCGCGCTCCCGATTCGACGCGATTCGCTTCTATCGCTTCTTGA 781
 QY 4509 CGAGTTTCTTGAGCGGAGTCT 4531
 Db 782 CGAGTTTCTTGAGCGGAGTCT 804
 RESULT 9
 CK288185
 LOCUS
 DEFINITION EST750907 Nicotiana benthamiana mixed tissue cDNA library,
 normalized, full-length Nicotiana benthamiana cDNA clone NEMB61 5',
 end, mRNA sequence.
 ACCSSION CK288185
 VERSION CK288185.1 GI:39865462
 KEYWORDS EST.
 SOURCE Nicotiana benthamiana
 ORGANISM Nicotiana benthamiana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 asterids; lamids; Solanales; Solanaceae; Nicotiana.
 REFERENCE 1 (bases 1 to 811)
 AUTHORS Buell,C.R., Hart,A., Zismann,V., Karamycheva,S.A., Day,B.,
 Staskawicz,B., Jin,H. and Baker,B.
 TITLE Generation of EST sequences from Nicotiana benthamiana
 JOURNAL Unpublished (2003)
 COMMENT Contact: Robin Buell
 The Institute for Genomic Research
 9712 Medical Center Dr, Rockville, MD 20850, USA
 Email: potato-array@tigr.org
 Clones can be requested from the University of Arizona Genomics
 Institute via <http://genome.arizona.edu/orders/>
 Seq primer: ATT TAG GTG ACA CTA TAG.
 FEATURES
 Location/Qualifiers
 1..811
 /organism="Nicotiana benthamiana"
 /mol_type="mRNA"
 /db_xref="taxon:4100"
 /clone="NEMB61"
 /tissue_type="abiotic and biotic stress-treated leaves,
 callus tissue and root tissue"
 /lab_host="DH10B-Tona"
 /clone_lib="Nicotiana benthamiana mixed tissue cDNA
 library, normalized, full-length"


```

Db 361 TTGAGAGAGCTGTGATGGAACGATGAGCAACATGGAAGCAGATCCAGTATCTTTTCAG 420
Qy 1456 ATAATGAAGCCAAATCTCTAGATGCTAAGAAATTTCCAAATTTCCAGATAACAATGATC 1515
Db 421 ATAATGAAGCCAAATCTCTAGATGCTAAGAAATTTCCAAATTTCCAGATAACAATGATC 480
Qy 1516 AAAGATTAAATGATGTTCTTTTCCAGCTAAATTTCTTACTTTCTTCCATCCAGAGACATG 1575
Db 481 AAAGATTAAATGATGTTCTTTTCCAGCTAAATTTCTTACTTTCTTCCATCCAGAGACATG 540
Qy 1576 AGAATATCATAGGGGATATCTCCAAGTCAATAGTAGGTCTGAACACCAACAGTACTTGATT 1635
Db 541 AGAATATCATAGGGGATATCTCCAAGTCAATAGTAGGTCTGAACACCAACAGTACTTGATT 600
Qy 1636 TGCAGTTCAGTATTGAACACATGAAATGCGAGATCCAAAGAGAAATGCAATTTAAACAACAG 1695
Db 601 TGCAGTTCAGTATTGAACACATGAAATGCGAGATCCAAAGAGAAATGCAATTTAAACAACAG 660
Qy 1696 AGGAGATCGTAAATAGAGAGCGTATATACAATGCATCAGCAAAATTAAGTCTCTAG 1755
Db 661 AGGAGATCGTAAATAGAGAGCGTATATACAATGCATCAGCAAAATTAAGTCTCTAG 720
Qy 1756 ATGAAAAACAAGTATATTGGAACAGGAAATATAAGGGGAAATGAAACTGTTGAATAATA 1815
Db 721 ATGAAAAACAAGTATATTGGAACAGGAAATATAAGGGGAAATGAAACTGTTGAATAATA 780
Qy 1816 TCACTAATGATCT 1828
Db 781 TCACTAATGATCT 793

```

RESULT 11

DR696294

LOCUS

DEFINITION SMV3-0053 Atlantic Salmon macrophage - Aeromonas in vivo infection
Salmo salar cDNA clone SMV3-0053, mRNA sequence.

ACCESSION

DR696294

VERSION

EST.

KEYWORDS

SOURCE

ORGANISM

Salmo salar

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Actinopterygii; Neopterygii; Teleostei; Euteleostei;

Protacanthopterygii; Salmoniformes; Salmonidae; Salmo;

1 (bases 1 to 767)

AUTHORS

Douglas, S.E., Tsou, S.C.M., Penny, S., Melville, K., Liebscher, R. and

Ewart, K.V.

TITLE

Unpublished (2005), Douglas, S.E., et al

JOURNAL

Unpublished (2005)

COMMENT

Contact: Douglas S

Genome Sciences

NRC Institute for Marine Biosciences

1411 Oxford St., Halifax, Nova Scotia B3H3Z1 Canada

Tel: (902) 426-4991

Fax: (902) 426-9413

Email: susan.douglas@nrc.ca

Seq primer: M13R.

Location/Qualifiers

1..767

/organism="Salmo salar"

/mol_type="mRNA"

/db_xref="taxon:8030"

/clone="SMV3-0053"

/cell_type="enriched macrophages"

/dev_stage="adult"

/clone_lib="Atlantic Salmon macrophage - Aeromonas in vivo

infection"

/note="Vector: pT-Adv; Libraries (SMV2, SMV3, SMV6, SMV7)

made by forward and reverse suppression subtractive

hybridization using salmon macrophage infected in vitro

for 0.5h (SMV2 and SMV3) or 2h (SMV6 and SMV7) with

Aeromonas salmonicida cultured in the intraperitoneal

space of fish vs. uninfected macrophages"

FEATURES

source

RESULT 12

CF409555

LOCUS

DEFINITION

CH3#061_D06MF Canine heart normalized cDNA Library in pluescript

Canis familiaris cDNA clone CH3#061_D06 5', mRNA sequence.

ACCESSION

CF409555

VERSION

EST.

KEYWORDS

SOURCE

ORGANISM

Canis familiaris (dog)

Canis familiaris

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;

Canis.

REFERENCE

1 (bases 1 to 949)

ORIGIN

Query Match 14.2%; Score 737.4; DB 8; Length 767;
Best Local Similarity 99.9%; Pred. No. 2e-190;
Matches 738; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

Qy 3783 ATTCCGGTATGACTGGGACACACAGACAATCGGCTGCTCTGATCGCGCGGTTCGGCT 3842
Db 12 ATTCCGGTATGACTGGGACACACAGACAATCGGCTGCTCTGATCGCGCGGTTCGGCT 71
Qy 3843 GTACAGCGAGGGCGCGCGGTTCCTTTTGTCAAGACCGACCTGTCCGTGCGCTGAATGA 3902
Db 72 GTACAGCGAGGGCGCGCGGTTCCTTTTGTCAAGACCGACCTGTCCGTGCGCTGAATGA 131
Qy 3903 ACTCAGAGCAGGACGCGGGCTATCGTGGCTGGCCACGACGCGGGCTTCCTTCGCGAGC 3962
Db 132 ACTCAGAGCAGGACGCGGGCTATCGTGGCTGGCCACGACGCGGGCTTCCTTCGCGAGC 191
Qy 3963 TGTGCTCGAGCTTGTCACTGAGCGGGAAGGGAAGTGGCTGCTATTTGGCGGAAGTGC 4022
Db 192 TGTGCTCGAGCTTGTCACTGAGCGGGAAGGGAAGTGGCTGCTATTTGGCGGAAGTGC 251
Qy 4023 GCAGGATCTCTGTCATCTCACCTTCTCTGCGGAGAAAGTATCCATCATGGCTGATGC 4082
Db 252 GCAGGATCTCTGTCATCTCACCTTCTCTGCGGAGAAAGTATCCATCATGGCTGATGC 311
Qy 4083 AATCGCGGGCTGCATACGCTTGTATCGGGCTTACCTGCGCCATTTCGACCAACGAGGAAACA 4142
Db 312 AATCGCGGGCTGCATACGCTTGTATCGGGCTTACCTGCGCCATTTCGACCAACGAGGAAACA 371
Qy 4143 TCGCATCGAGCGACACGCTACTCGGATGGAAGCGCGTCTTGTGATCAGGATGATCTGGA 4202
Db 372 TCGCATCGAGCGACACGCTACTCGGATGGAAGCGCGTCTTGTGATCAGGATGATCTGGA 431
Qy 4203 CGAAGACATCAGGGGCTCGCGCCAGCGCACTGTTCCGACAGGCTCAAGGCGCGCATGCC 4262
Db 432 CGAAGACATCAGGGGCTCGCGCCAGCGCACTGTTCCGACAGGCTCAAGGCGCGCATGCC 491
Qy 4263 CGACGGCGAGGATCTCGTGTGACCCATGCGATGCGCTGCTTGGCCGAATATCATGGTGA 4322
Db 492 CGACGGCGAGGATCTCGTGTGACCCATGCGATGCGCTGCTTGGCCGAATATCATGGTGA 551
Qy 4323 AATGGCGGCTTTTCTGATTCATCGACTGTGGCGCGCTGGGTGTGGCGGACCGCTATCA 4382
Db 552 AATGGCGGCTTTTCTGATTCATCGACTGTGGCGCGCTGGGTGTGGCGGACCGCTATCA 611
Qy 4383 GGACATAGCTTGGCTACCGCTGATATGCTGAAGAGCTTGGCGGCGAATGGGCTGACCG 4442
Db 612 GGACATAGCTTGGCTACCGCTGATATGCTGAAGAGCTTGGCGGCGAATGGGCTGACCG 671
Qy 4443 CTTCTCGTCTTTTACCGTATCGCGCTCCCGATTCGAGCGCATTCGCTTCTATCGCT 4502
Db 672 CTTCTCGTCTTTTACCGTATCGCGCTCCCGATTCGAGCGCATTCGCTTCTATCGCT 731
Qy 4503 TCTTACGAGTCTTCTTGA 4521
Db 732 TCTTACGAGTCTTCTTGA 750

```

CF409555 949 bp mRNA linear EST 02-SEP-2003
CH3#061_D06MF Canine heart normalized cDNA Library in pluescript
Canis familiaris cDNA clone CH3#061_D06 5', mRNA sequence.

GI:34410801

Canis familiaris (dog)

Canis familiaris

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;

Canis.

REFERENCE

1 (bases 1 to 949)

3546 ATGGCGATAGCTAGCTGGCGGTTTATGACACAGCAGCAGCGAATGCGAGCTGG 3605
 Db ATGGCGATAGCTAGCTGGCGG-TTTATGTCAGCAGCAGCATCCGATTTGCCAGCTGG 646
 3606 GCGCGCCTCTGTAAGTGTGGGAAGCCCTGCAAAAGTAACTGGATGGCTTCTTGGCGGCC 3665
 Db GCGCGCCTTCTGTAAGTGTGGGAAGCCCTGCAAAAGTAACTGGATGGCTTCTTGGCGGCC 586
 3666 AAGGATCTGATGGCGCAGGGGATCAAGATCTGATCAAGAGACAGATGAGATCGTTTCG 3725
 Db AAGGATCTGATGGCGCAGGGGATCAAGCTCTGATCAAGAGACAGATGAGATCGTTTCG 526
 3726 CATGATTGAACAAGTGTGACGACGAGTCTCCGCGCGCTTGGGTGGAGAGCTATT 3785
 Db CATGATTGAACAAGTGTGACGACGAGTCTCCGCGCGCTTGGGTGGAGAGCTATT 466
 3786 CGGCTATGACTGGGCACACAGACAAATCGGCTGTCTGATCCGCGCTTCCGGCTGTC 3845
 Db CGGCTATGACTGGGCACACAGACAAATCGGCTGTCTGATCCGCGCTTCCGGCTGTC 406
 3846 AGCGCAGGGGCGCGGTTCTTTTGTCAAGACCGACCTGTCCGGTGCCTTGAATGAAT 3905
 Db AGCGCAGGGGCGCGGTTCTTTTGTCAAGACCGACCTGTCCGGTGCCTTGAATGAAT 346
 3906 CGAGACGAGCAGCGCGGCTATCGTGTGCGCACAGCGGGCGTCTTCCGCGAGCTGT 3965
 Db CGAGACGAGCAGCGCGGCTATCGTGTGCGCACAGCGGGCGTCTTCCGCGAGCTGT 286
 3966 GCTCGACGTTCTACTGAAGCGGGAAGGACTGG-CTGCTATTGGCGGAGTGGCGGGGC 4024
 Db GCTCGACGTTCTACTGAAGCGGGAAGGACTGGCTGCTGTTTGGCGGAGTGGCGGGGC 226
 4025 AGGATCTCTGTCTATCTCACTTGTCTGCTCGGAGAAAGTATCCATCATGCTCATGCAA 4084
 Db AGGATCTCTGTCTATCTCACTTGTCTGCTCGGAGAAAGTATCCATCATGCTCATGCAA 166
 4085 TGGCGCGCTGATACGTTGATCGGCTACTGCGCATTCGACCAACAGCGAAACATC 4144
 Db TGGCGCGCTGATACGTTGATCGGCTACTGCGCATTCGACCAACAGCGAAACATC 106
 4145 GCATCGAGCAGCAGCTACTCGATGGAAGCGGTC-TTGTGCTAGCAGGATGATC-TGGA 4202
 Db GCATCGAGCAGCAGCTACTCGATGGAAGCGGTCATTGTGCTAGCAGGATGATCGTGA 46
 4203 CGAAGAGCATCAGGGGCTCGCGCAGCGAAGTGTTCGCCAGGCT 4247
 Db CGAAGAGCATCAGGGGCTCGCGCAGCGAAGTGTTCGCCAGGCT 1

RESULT 14

CK288711
 LOCUS EST751433 Nicotiana benthamiana 878 bp mRNA linear EST 02-AUG-2004
 DEFINITION normalized, full-length Nicotiana benthamiana cDNA clone NEMB149 5',
 end, mRNA sequence.
 CK288711
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

Nicotiana benthamiana
 Nicotiana benthamiana
 Sukariota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 asterids; lamids; Solanales; Solanaceae; Nicotiana.
 1. (bases 1 to 878)

Buell.C.R., Hart, A., Zismann, V., Karamycheva, S.A., Day, B.,
 Staskawicz, B., Jin, H. and Baker, B.
 Generation of EST sequences from Nicotiana benthamiana
 Unpublished (2003)
 CONTACT: Robin Buell

The Institute for Genomic Research
 9712 Medical Center Dr, Rockville, MD 20850, USA
 Email: potato-array@tigr.org
 Clones can be requested from the University of Arizona Genomics

Institute via <http://genome.arizona.edu/orders/>
 Seq primer: ATT TAG GTG ACA CTA TAG.

FEATURES

Source

1. 878
 Location/Qualifiers
 /organism="Nicotiana benthamiana"
 /mol_type="mRNA"
 /db_xref="taxon:4100"
 /clone="NEMB149"
 /tissue_type="abiotic and biotic stress-treated leaves,
 callus tissue and root tissue"
 /lab_host="DH10B-TonA"
 /clone_lib="Nicotiana benthamiana mixed tissue cDNA
 library, normalized, full-length"
 /note="vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;
 supplier: RNA was isolated from Nicotiana benthamiana
 tissues that include callus, roots from liquid culture
 grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr),
 cold-stressed leaves (5 C 3 hr, 6hr), and pathogen
 challenged leaves (Pseudomonas syringae pv tomato 12 hr;
 Xanthomonas campestris pv campestris 12 hr, 18hr;
 Pseudomonas syringae pv phaseolicola 18hr, and Xanthomonas
 campestris pv vesicatoria 18hr). RNA was isolated from
 these tissues and pooled in approximately equal molar
 amounts."

ORIGIN

Query Match 11.7%; Score 605.4; DB 7; Length 878;
 Best Local Similarity 99.8%; Pred. No. 4e-154;
 Matches 606; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 3936 GGCCACGACGGGGTTCCTTGGCGAGCTGTCTCGAGCTGTCTCACTGAGCGGGAAGGGA 3995
 Db 1 GGCCACGACGGGGTTCCTTGGCGAGCTGTCTCGAGCTGTCTCACTGAGCGGGAAGGGA 60
 Qy 3996 CTGCTCTCTATTGGGGCAAGTCCGGGGAGGATCTCTGTCATCTCACCTTGTCTCTGC 4055
 Db 61 CTGCTCTCTATTGGGGCAAGTCCGGGGAGGATCTCTGTCATCTCACCTTGTCTCTGC 120
 Qy 4056 CGAGAAAGTATCCATCATATGTCATGCAATGCGGGCTGTCATACGCTTGCATCGGCTAC 4115
 Db 121 CGAGAAAGTATCCATCATATGTCATGCAATGCGGGCTGTCATACGCTTGCATCGGCTAC 180
 Qy 4116 CTGCCCCATTCGACCAACGCGAAACATTCGCATCGAGCGAGCAGCTACTCGGATGGAAGC 4175
 Db 181 CTGCCCATTCGACCAACGCGAAACATTCGCATCGAGCGAGCAGCTACTCGGATGGAAGC 240
 Qy 4176 CGGCTTGTGATCAGATGATCTGGACGAGAGCATCAGGGGCTCGCGCCAGCGGACT 4235
 Db 241 CGGCTTGTGATCAGATGATCTGGACGAGAGCATCAGGGGCTCGCGCCAGCGGACT 300
 Qy 4236 GTTCGCGAGCTCAAGCGCGCATGCGCGAGGAGTCTCGTGTGACCCATGGCGA 4295
 Db 301 GTTCGCGAGCTCAAGCGCGCATGCGCGAGGAGTCTCGTGTGACCCATGGCGA 360
 Qy 4296 TGCTCTGCTCCGAATATCATGTGGAAATAGCCCGCTTTTCTGGATTCATTCGACTGTGG 4355
 Db 361 TGCTCTGCTCCGAATATCATGTGGAAATAGCCCGCTTTTCTGGATTCATTCGACTGTGG 420
 Qy 4356 CCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCTTGGCTACCGTGTATTTGCTGA 4415
 Db 421 CCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCTTGGCTACCGTGTATTTGCTGA 480
 Qy 4416 AGAGCTTGGCGGCAATGGGCTGACCGCTTCTCGTCTTTTACGGTATCGCGCTCCCGA 4475
 Db 481 AGAGCTTGGCGGCAATGGGCTGACCGCTTCTCGTCTTTTACGGTATCGCGCTCCCGA 540
 Qy 4476 TTGCGACGCAATCGCTTCTATCGCTTCTTGAACGAGTCTTCTTGAACGCGGACTCTGGGG 4535
 Db 541 TTGCGACGCAATCGCTTCTATCGCTTCTTGAACGAGTCTTCTTGAACGCGGACTCTGGGG 600
 Qy 4536 TTGCGATA 4542
 Db 601 TTGCGATA 607

| | | | | | |
|-----------------------|---|--|-----------|-------------|-----------------|
| RESULT 15 | DR696687 | 591 bp | mRNA | linear | EST 13-JUL-2005 |
| LOCUS | SMV7-0178 | Atlantic Salmon macrophage - Aeromonas in vivo infection | | | |
| DEFINITION | Salmo salar | CDNA clone SMV7-0178, mRNA sequence. | | | |
| ACCESSION | DR696687 | | | | |
| VERSION | DR696687.1 | GI:70787047 | | | |
| KEYWORDS | EST. | | | | |
| SOURCE | Salmo salar | (Atlantic salmon) | | | |
| ORGANISM | Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Protacanthopterygii; Salmoniformes; Salmonidae; Salmo. | | | | |
| REFERENCE | 1 | (bases 1 to 591) | | | |
| AUTHORS | Douglas,S.E., Tsou,S.C.M., Penny,S., Melville,K., Liebscher,R. and Ewart,K.V. | | | | |
| TITLE | Unpublished (2005), Douglas,S.E., et al | | | | |
| JOURNAL | Unpublished (2005) | | | | |
| COMMENT | Contact: Douglas S Genome Sciences NRC Institute for Marine Biosciences 1411 Oxford St., Halifax, Nova Scotia B3H3Z1 Canada Tel: (902) 426-4991 Fax: (902) 426-9413 Email: susan.douglas@nrc.ca Seq primer: M13p. | | | | |
| FEATURES | Location/Qualifiers | | | | |
| source | 1..591 | | | | |
| | /organism="Salmo salar" | | | | |
| | /mol_type="mRNA" | | | | |
| | /db_xref="taxon:8030" | | | | |
| | /clone="SMV7-0178" | | | | |
| | /call_type="enriched macrophages" | | | | |
| | /dev_stage="adult" | | | | |
| | /clone_lib="Atlantic Salmon macrophage - Aeromonas in vivo infection" | | | | |
| | /note="vector: pT-Adv; Libraries (SMV2, SMV3, SMV6, SMV7) made by forward and reverse suppression subtractive hybridization using salmon macrophage infected in vitro for 0.5h (SMV2 and SMV3) or 2h (SMV6 and SMV7) with Aeromonas salmonicida cultured in the intraperitoneal space of fish vs. uninfected macrophages" | | | | |
| ORIGIN | | | | | |
| Query Match | 11.4% | Score 589.4; | DB 8; | Length 591; | |
| Best Local Similarity | 99.8% | Pred. No. 8.7e-150; | | | |
| Matches 590; | Conservative 0; | Mismatches 1; | Indels 0; | Gaps 0; | |
| QY | 3812 | TCGGCTGCTCTGATGCGCGCGTGTTCGGGTGTTCAGCGAGGCGCCCGTCTCTTTTG | 3871 | | |
| DB | 1 | TCGGCTGCTCTGATGCGCGCGTGTTCGGGTGTTCAGCGAGGCGCCCGTCTCTTTTG | 60 | | |
| QY | 3872 | TCAAGACCGACCTGTTCGGTGCCTGAATGAATGCAGAGCAGGCGCGGCTATCGT | 3931 | | |
| DB | 61 | TCAAGACCGACCTGTTCGGTGCCTGAATGAATGCAGAGCAGGCGCGGCTATCGT | 120 | | |
| QY | 3932 | GGCTGGGCCACGACGGGCGTTCTTTGCGCAGCTGTGCTCGACGTTGTCACTGAACGGGAA | 3991 | | |
| DB | 121 | GGCTGGGCCACGACGGGCGTTCTTTGCGCAGCTGTGCTCGACGTTGTCACTGAACGGGAA | 180 | | |
| QY | 3992 | GGGACTGGCTGTATTTGGGGAAAGTGGCGGGCAGGATCTCTGTATCTCACTTCGTC | 4051 | | |
| DB | 181 | GGGACTGGCTGTATTTGGGGAAAGTGGCGGGCAGGATCTCTGTCACTCCACCTTCGTC | 240 | | |
| QY | 4052 | CTGCCGAGAAGTATCCATCATGCTCGATGCAATGCGCGGCTGCATACGCTTCATCCG | 4111 | | |
| DB | 241 | CTGCCGAGAAGTATCCATCATGCTCGATGCAATGCGCGGCTGCATACGCTTCATCCG | 300 | | |
| QY | 4112 | CTACCTGCCCATTCGACCAACGCGAAACATCGCATCGACGCGAGCACTGATCCGATGG | 4171 | | |
| DB | 301 | CTACCTGCCCATTCGACCAACGCGAAACATCGCATCGACGCGAGCACTGATCCGATGG | 360 | | |

GenCore version 5.1.7

Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 27, 2006, 08:16:10 ; Search time 2790 Seconds
(without alignments)
12366.707 Million cell updates/sec

Title: US-10-618-570-1

Perfect score: 5177

Sequence: 1 ttgaagaccaccaccgta.....agcgggggttttctattgg 5177

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_21.*

1: Geneseqn1980s.*

2: Geneseqn1990s.*

3: Geneseqn2000s.*

4: Geneseqn2001as.*

5: Geneseqn2001bs.*

6: Geneseqn2002as.*

7: Geneseqn2002bs.*

8: Geneseqn2003as.*

9: Geneseqn2003bs.*

10: Geneseqn2003cs.*

11: Geneseqn2003ds.*

12: Geneseqn2004as.*

13: Geneseqn2004bs.*

14: Geneseqn2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|----|--------------------|
| 1 | 5165.8 | 99.8 | 5177 | 2 | Aaz09997 Bovine sc |
| 2 | 2077.2 | 40.1 | 7699 | 2 | Aav81410 Plasmid p |
| 3 | 1972.6 | 38.1 | 7436 | 13 | Adq0675 Bcl-XL ex |
| 4 | 1916.6 | 37.0 | 7546 | 2 | Aax90485 Plasmid r |
| 5 | 1661.6 | 32.1 | 8316 | 2 | Aac74450 myod retr |
| 6 | 1644.6 | 31.8 | 7165 | 2 | Aax90483 Plasmid r |
| 7 | 1629.4 | 31.5 | 5594 | 2 | Aav33629 GENSA 981 |
| 8 | 1588.6 | 30.7 | 5130 | 6 | Adm68972 LSRNL vec |
| 9 | 1588.6 | 30.7 | 5130 | 12 | Adm68972 LSRNL vec |
| 10 | 1588.6 | 30.7 | 5130 | 14 | Adw28107 Nucleotid |
| 11 | 1588.6 | 30.7 | 5130 | 14 | Adw28107 Nucleotid |
| 12 | 1486 | 28.7 | 5874 | 2 | Aax90484 Plasmid r |
| 13 | 1486 | 28.7 | 6141 | 2 | Aax90481 Plasmid r |
| 14 | 1486 | 28.7 | 6522 | 2 | Aax90482 Plasmid r |
| 15 | 1486 | 28.7 | 6620 | 6 | Adc82653 Plasmid v |
| 16 | 1486 | 28.7 | 9115 | 2 | Aac74449 myod retr |
| 17 | 1484.4 | 28.7 | 6365 | 2 | Aaq41173 Plasmid L |
| 18 | 1481 | 28.6 | 6046 | 12 | Adg83262 Clone pLX |
| 19 | 1463 | 28.3 | 6825 | 14 | Aeb48081 pPB-ctAP- |

| | | | | | |
|----|--------|------|------|----|--------------------|
| 20 | 1463 | 28.3 | 6851 | 14 | AEB48080 |
| 21 | 1453.6 | 28.1 | 5865 | 2 | AAV04002 Retrovira |
| 22 | 1448.4 | 28.0 | 8591 | 2 | Aat84562 Plasmid p |
| 23 | 1448.4 | 28.0 | 8591 | 2 | Aat84561 Plasmid p |
| 24 | 1448.4 | 28.0 | 8591 | 2 | Aat87083 Plasmid p |
| 25 | 1448.4 | 28.0 | 8591 | 2 | Aat87084 Plasmid p |
| 26 | 1448.4 | 28.0 | 8591 | 2 | AAV04866 CDNA enco |
| 27 | 1448.4 | 28.0 | 8591 | 2 | AAV04865 CDNA enco |
| 28 | 1448.4 | 28.0 | 8591 | 2 | AAV05850 APP-REP 7 |
| 29 | 1448.4 | 28.0 | 8591 | 2 | AAV05849 APP-REP 7 |
| 30 | 1448 | 28.0 | 6444 | 6 | Aai67595 Nucleotid |
| 31 | 1441 | 27.8 | 7170 | 6 | Adp28272 LSRNL vec |
| 32 | 1434 | 27.7 | 6700 | 4 | Aah76187 Nucleotid |
| 33 | 1434 | 27.7 | 8518 | 4 | Aah74524 Nucleotid |
| 34 | 1434 | 27.7 | 8518 | 4 | Aah76188 Nucleotid |
| 35 | 1432.4 | 27.7 | 6700 | 4 | Aah74523 Nucleotid |
| 36 | 1429.8 | 27.6 | 8078 | 10 | AdB52213 |
| 37 | 1425 | 27.5 | 5689 | 2 | Aaz11445 Retrovira |
| 38 | 1425 | 27.5 | 5689 | 2 | Aax61061 Retrovira |
| 39 | 1425 | 27.5 | 5689 | 3 | AAx96208 Nucleotid |
| 40 | 1425 | 27.5 | 9849 | 3 | AAx96221 Nucleotid |
| 41 | 1352 | 26.1 | 7925 | 12 | ADP74702 |
| 42 | 1292.8 | 25.0 | 9830 | 3 | AAx96220 Nucleotid |
| 43 | 1206.8 | 23.3 | 6505 | 9 | ADA12886 Murine MS |
| 44 | 1186.6 | 22.9 | 7185 | 12 | ADP74699 Novel bic |
| 45 | 1186.6 | 22.9 | 7185 | 12 | ADP74701 Novel bic |

ALIGNMENTS

RESULT 1

AAZ09997

ID AAZ09997 standard; cDNA; 5177 BP.

XX AC AAZ09997;

XX AC AAZ09997;

DT 03-DEC-1999 (first entry)

DE Bovine scavenger receptor class A (ScR)/avidin fusion protein cDNA.

XX Scavenger receptor class A; ScR; avidin; fusion protein; bovine; ECD;

KW membrane-spanning domain; extracellular domain; biotin-binding activity;

KW endocytosis; ss.

XX Synthetic.

OS Bos taurus.

XX Bos taurus.

FH Key Location/Qualifiers

FT CDS 1071..2269

FT CDS /*tag= a

FT CDS /product= "ScR/avidin fusion protein"

XX WO9942577-A2.

XX 26-AUG-1999.

XX 23-FEB-1999; 99WO-GB000546.

XX 23-FEB-1998; 98GB-00003757.

XX 24-JUN-1998; 98GB-00013653.

XX (EURO-) EUROGENE LTD.

XX Yla-Herttua S, Kulomaa M, Lehtolainen P, Marjomaki V, Airenne K;

XX WPI; 1999-561345/47.

XX P-PSDB; AAY03098.

XX New fusion proteins having an extracellular domain with biotin-binding

XX activity, used to target biotinylated molecules to specific sites in

XX tissues.

XX

QY 1921 ACGATCTGGGCTCCAAATGACCATCGGGGCTGTGAACAGCAGAGGTGAATTCACAGGCA 1980
Db 1921 ACGATCTGGGCTCCAAATGACCATCGGGGCTGTGAACAGCAGAGGTGAATTCACAGGCA 1980
QY 1981 CCTACATCACAGCGGTAAACAGCCACATCAAAATGAGATCAAAAGAGTCAACACTGCATGSGA 2040
Db 1981 CCTACATCACAGCGGTAAACAGCCACATCAAAATGAGATCAAAAGAGTCAACACTGCATGSGA 2040
QY 2041 CACAAAAACACATCAAAAGAGGACCCAGCCACCTTTGGCTTCAACCGTCAATTTGGAAGT 2100
Db 2041 CACAAAAACACATCAAAAGAGGACCCAGCCACCTTTGGCTTCAACCGTCAATTTGGAAGT 2100
QY 2101 TTTTCAGAGTCACACACTGTCTTTCAGGGCCAGTGTTCATAGACAGGAAATGGGAAGG 2160
Db 2101 TTTTCAGAGTCACACACTGTCTTTCAGGGCCAGTGTTCATAGACAGGAAATGGGAAGG 2160
QY 2161 TCCTGAAAGACCATGTGGCTGTGGCTCAAGTGTAAATGACATTTGGTGTGATGCTGGAAG 2220
Db 2161 TCCTGAAAGACCATGTGGCTGTGGCTCAAGTGTAAATGACATTTGGTGTGATGCTGGAAG 2220
QY 2221 CTACACAGGTCGGCATCAACATCTTCACTCGCCTCGGCACACAGAGAGTGAAGTG 2280
Db 2221 CTACACAGGTCGGCATCAACATCTTCACTCGCCTCGGCACACAGAGAGTGAAGTG 2280
QY 2281 ACCAAGGTCCTCTGGAATCCAGGTGAAGAGGAGATAGAGCCCTCTCGGACAAAATGG 2340
Db 2281 ACCAAGGTCCTCTGGAATCCAGGTGAAGAGGAGATAGAGCCCTCTCGGACAAAATGG 2340
QY 2341 TATACCAAGCTTCCAGGTCCTAAATAGGTACTCCAGGTCTTAAAGGTGATCGGGGGATCT 2400
Db 2341 TATACCAAGCTTCCAGGTCCTAAATAGGTACTCCAGGTCTTAAAGGTGATCGGGGGATCT 2400
QY 2401 CTGGTTTACCTGGAGTTCAGGATTTCCAGGACCAATGGGGAGAACCGGGAGCCAGGAC 2460
Db 2401 CTGGTTTACCTGGAGTTCAGGATTTCCAGGACCAATGGGGAGAACCGGGAGCCAGGAC 2460
QY 2461 TTAATGGACAAAAGGCGCAGAAAGGAGGAGGAGGAGGAGCATGCAAAAGACAATCTA 2520
Db 2461 TTAATGGACAAAAGGCGCAGAAAGGAGGAGGAGGAGGAGCATGCAAAAGACAATCTA 2520
QY 2521 ATACAGTCCGACTGGTGGGCGAGCGCCCTCAGAGGAGGAGTGGAGATTTTTCACG 2580
Db 2521 ATACAGTCCGACTGGTGGGCGAGCGCCCTCAGAGGAGGAGTGGAGATTTTTCACG 2580
QY 2581 AAGGCCAGTGGGTACGGTGTGTGACGACCCCTCGGAACTCGTGGAGGACTGCTGCT 2640
Db 2581 AAGGCCAGTGGGTACGGTGTGTGACGACCCCTCGGAACTCGTGGAGGACTGCTGCT 2640
QY 2641 GCAGAGCTTGGGATACAAAGGTGTTCAAAGTGTGCAATAGCCGAGCTTATTTTGGAAAAG 2700
Db 2641 GCAGAGCTTGGGATACAAAGGTGTTCAAAGTGTGCAATAGCCGAGCTTATTTTGGAAAAG 2700
QY 2701 GTACGGGTCCAAATATGGCTGAATGAATATTTTGTTCGGGAAAGAGTCATCCATTTGAG 2760
Db 2701 GTACGGGTCCAAATATGGCTGAATGAATATTTTGTTCGGGAAAGAGTCATCCATTTGAG 2760
QY 2761 AGTCAGAAATTAGACAGTGGGCTGTGAGAGCTGTTCGACGACCAAGAGTCTCGGGGTC 2820
Db 2761 AGTCAGAAATTAGACAGTGGGCTGTGAGAGCTGTTCGACGACCAAGAGTCTCGGGGTC 2820
QY 2821 ACTTTGCACTACATAATATGATATTTTCAATTCACATTTTAAACCTGTTATAAAGTG 2880
Db 2821 ACTTTGCACTACATAATATGATATTTTCAATTCACATTTTAAACCTGTTATAAAGTG 2880
QY 2881 ATTTTTCCTTTCCTTCACTAAATAGCTTAAATATTTTAAAGAACTTAAGAAATTTT 2940
Db 2881 ATTTTTCCTTTCCTTCACTAAATAGCTTAAATATTTTAAAGAACTTAAGAAATTTT 2940
QY 2941 ATCCACAGAAAGGAATATTTTAAATACCTTGGATATAACATATAAATAGCTTCATATTT 3000
Db 2941 ATCCACAGAAAGGAATATTTTAAATACCTTGGATATAAATAGCTTCATATTT 3000
QY 3001 GCTTCAAATACCAGAACCATTTTCACTTCTAGGTTTTTAAAGTGGCTCGTGGCGAAATG 3060

Db 3001 GCTTCAAATACCAGAACCATTTTCACTTCTAGGTTTTTAAAGTGGCTCGTGGCGAAATG 3060
QY 3061 ATCCCTCTCAGGATATAGTATTTTCGCTTTTGCATAGGAGGGGAAATAGTAGTCTTATGC 3120
Db 3061 ATCCCTCTCAGGATATAGTATTTTCGCTTTTGCATAGGAGGGGAAATAGTAGTCTTATGC 3120
QY 3121 AATACTCTTGTAGTCTTGGCAACATGTTAAACATGATGATTAGCAACATGCTTTACAGAGA 3180
Db 3121 AATACTCTTGTAGTCTTGGCAACATGTTAAACATGATGATTAGCAACATGCTTTACAGAGA 3180
QY 3181 GAAAAACGACCGCTCAGCCGATTTGGTGAAGTAAAGTGTGATCGATCGTGTCTTATTAGG 3240
Db 3181 GAAAAACGACCGTGCATGCGGATTTGGTGAAGTAAAGTGTGATCGATCGTGTCTTATTAGG 3240
QY 3241 AAGCAACACAGACGGGTCTGACATGGAATGGACGAACCACTGAATTTCCGCAATTCAGAGAT 3300
Db 3241 AAGCAACACAGACGGGTCTGACATGGAATGGACGAACCACTGAATTTCCGCAATTCAGAGAT 3300
QY 3301 ATTGTATTTAAGTCTAGCTCGATACAGCAACCGCATTTGACCATTTACCAATTTGGT 3360
Db 3301 ATTGTATTTAAGTCTAGCTCGATACAGCAACCGCATTTGACCATTTACCAATTTGGT 3360
QY 3361 GTGCACCTCCAGCTTCAAGCTTCCAGTTCGCGCAGCACTCAGGCGCAGAGGCTGCTAAAGGAG 3420
Db 3361 GTGCACCTCCAGCTTCAAGCTTCCAGTTCGCGCAGCACTCAGGCGCAGAGGCTGCTAAAGGAG 3420
QY 3421 CGGAAACAGCTAGAAAGCAGTCCGAGAAACGGTGTGACCCCGATGAATGTCCAGCTAC 3480
Db 3421 CGGAAACAGCTAGAAAGCAGTCCGAGAAACGGTGTGACCCCGATGAATGTCCAGCTAC 3480
QY 3481 TGGGCTATCTGGACAAAGGAAAACGCAAGCGCAAGAGAGAGTAGCTTTCAGTGGG 3540
Db 3481 TGGGCTATCTGGACAAAGGAAAACGCAAGCGCAAGAGAGAGTAGCTTTCAGTGGG 3540
QY 3541 CTTCATATGGCGATAGCTAGACTGGGCGGTTTTATAGGACAGCAAGCGGAATTTGCCA 3600
Db 3541 CTTCATATGGCGATAGCTAGACTGGGCGGTTTTATAGGACAGCAAGCGGAATTTGCCA 3600
QY 3601 GCTGGGCGGCTCTGCTGTAAGTGGGAGCCCTGCAAAAGTAAACTCGATGGCTTTCTTG 3660
Db 3601 GCTGGGCGGCTCTGCTGTAAGTGGGAGCCCTGCAAAAGTAAACTCGATGGCTTTCTTG 3660
QY 3661 CCGCAAGGATCTGATGGCGCAGGGATCAAGATCTGATCAAGAGACAGGATAGGATCG 3720
Db 3661 CCGCAAGGATCTGATGGCGCAGGGATCAAGATCTGATCAAGAGACAGGATAGGATCG 3720
QY 3721 TTTTCGATGATGAAACAGATGATTCAGCGAGGTTCTCGGCGGCTTGGGCTGGAGAGG 3780
Db 3721 TTTTCGATGATGAAACAGATGATTCAGCGAGGTTCTCGGCGGCTTGGGCTGGAGAGG 3780
QY 3781 CTATTCGGCTATGACTGGGCAACACAGCAATCGGCTGCTCTGATGCGCGCTGTTCCGG 3840
Db 3781 CTATTCGGCTATGACTGGGCAACACAGCAATCGGCTGCTCTGATGCGCGCTGTTCCGG 3840
QY 3841 CTGTACGCGAGGGGCGCCGCTTCTTTTGTCAAGACCGACCTGTCCGCTGCCCTGAAT 3900
Db 3841 CTGTACGCGAGGGGCGCCGCTTCTTTTGTCAAGACCGACCTGTCCGCTGCCCTGAAT 3900
QY 3901 GAACTGCGAGGACGAGCGCGCTATCGTGGCTGGCGACGACGGGCTTCTTTGGCGCA 3960
Db 3901 GAACTGCGAGGACGAGCGCGCTATCGTGGCTGGCGACGACGGGCTTCTTTGGCGCA 3960
QY 3961 GCTGTGCTCGACGTTGTCACTGAAGCGGAAAGGACTGGCTGCTATTTGGGCGAAGTCCG 4020
Db 3961 GCTGTGCTCGACGTTGTCACTGAAGCGGAAAGGACTGGCTGCTATTTGGGCGAAGTCCG 4020
QY 4021 GGGCAGGATCTCTGTCTCATCTCACCTTGTCTTCCCGAGAAAGTATCATCATGGCTGAT 4080
Db 4021 GGGCAGGATCTCTGTCTCATCTCACCTTGTCTTCCCGAGAAAGTATCATCATGGCTGAT 4080
QY 4081 GCAATGCGGCGGTGCATACGCTTGTGATCGGCTACCTGCCCATTTTCGACCAACAGCGGAA 4140
Db 4081 GCAATGCGGCGGTGCATACGCTTGTGATCGGCTACCTGCCCATTTTCGACCAACAGCGGAA 4140

Db 4081 GCAATGCGCGCTGCATACGCTTGATCCGGCTACCTGCCCATTCGACCACCAAGCGAAA 4140
Qy CATGCGATCGAGCGAGCAGCTACTCGGATGGAAGCGGCTGTGCGATCAGGATGATCTG 4200
Db CATGCGATCGAGCGAGCAGCTACTCGGATGGAAGCGGCTGTGCGATCAGGATGATCTG 4200
Qy GACGAAGAGCATCAGGGGCTCGCGCCAGCCGAACTGTTCGCGAGGCTCAAGCGCGCATG 4260
Db GACGAAGAGCATCAGGGGCTCGCGCCAGCCGAACTGTTCGCGAGGCTCAAGCGCGCATG 4260
Qy CCGAGCGCGAGGATCTCGTGTGACCCATGGGATGCTCTGTCGCGAATATCATGGTG 4320
Db CCGAGCGCGAGGATCTCGTGTGACCCATGGGATGCTCTGTCGCGAATATCATGGTG 4320
Qy GAAATGSCCGCTTTCTTGGAATTCATGACGTGCGCGGCTGGGTGGCGGACCGCTAT 4380
Db GAAATGSCCGCTTTCTTGGAATTCATGACGTGCGCGGCTGGGTGGCGGACCGCTAT 4380
Qy CAGGACATAGCGTTGGCTACCGTGATATTCTGAAGAGCTTTGGCGGCGAATGGCTGAC 4440
Db CAGGACATAGCGTTGGCTACCGTGATATTCTGAAGAGCTTTGGCGGCGAATGGCTGAC 4440
Qy CGCTTCCTCGCTTTACGGTATCGCGCTCCCGATTCCGAGCGCATCGCCTTCTATCGC 4500
Db CGCTTCCTCGCTTTACGGTATCGCGCTCCCGATTCCGAGCGCATCGCCTTCTATCGC 4500
Qy CTTCTTGACGAGTTCTTCTGAGCGGACTCTGGGGTTGATATAAATAAAGATTTTATTT 4560
Db CTTCTTGACGAGTTCTTCTGAGCGGACTCTGGGGTTGATATAAATAAAGATTTTATTT 4560
Qy AGTCTCCAGAAAAAGGGGGGAATGAAGACCCCACTGTAGTGTTCGCAAGCTAGCTTAA 4620
Db AGTCTCCAGAAAAAGGGGGGAATGAAGACCCCACTGTAGTGTTCGCAAGCTAGCTTAA 4620
Qy GTAAAGCCCATTTTGCAGGCTGGAATAATACATACTGAGAAATAGAGAAATTCAGATCA 4680
Db GTAAAGCCCATTTTGCAGGCTGGAATAATACATACTGAGAAATAGAGAAATTCAGATCA 4680
Qy AGGTGAGGAACAGATGGAACAGCTGAAATATGGGCCAAACAGGATATCTGTGTTAGCAGT 4740
Db AGGTGAGGAACAGATGGAACAGCTGAAATATGGGCCAAACAGGATATCTGTGTTAGCAGT 4740
Qy TCCTGCCCGGCTCAGGGCCAAAGACAGATGGAACAGCTGAAATATGGGCCAAACAGGATA 4800
Db TCCTGCCCGGCTCAGGGCCAAAGACAGATGGAACAGCTGAAATATGGGCCAAACAGGATA 4800
Qy TCTGTGTAAGCAGTTCTGCCCCGGCTCAGGGCCAAAGACAGATGTTGCCAGATCGG 4860
Db TCTGTGTAAGCAGTTCTGCCCCGGCTCAGGGCCAAAGACAGATGTTGCCAGATCGG 4860
Qy TCCAGCCCTCAGCAGTTCTTAGAGAACCATCAGATGTTCCAGGGTCCCCCAAGGACCTG 4920
Db TCCAGCCCTCAGCAGTTCTTAGAGAACCATCAGATGTTCCAGGGTCCCCCAAGGACCTG 4920
Qy AAATGACCTGTGCTTATTTGAATCAACCAATCAGTTCGCTTCTGTTTCGCGC 4980
Db AAATGACCTGTGCTTATTTGAATCAACCAATCAGTTCGCTTCTGTTTCGCGC 4980
Qy GCTTCTGTCTCCCGAGCTCAATAAAGAGCCCAACACCCCTCACTCGGGGCGCGAGTCT 5040
Db GCTTCTGTCTCCCGAGCTCAATAAAGAGCCCAACACCCCTCACTCGGGGCGCGAGTCT 5040
Qy CCGATTGACTAGTCGCGCGGGTACCGGTGATTCGAATAAACCTCTTTCGAGTTGCATCC 5100
Db CCGATTGACTAGTCGCGCGGGTACCGGTGATTCGAATAAACCTCTTTCGAGTTGCATCC 5100
Qy GACTTGTGCTCGCTGTCTTGGAGGGTCTCTCTGAGTGATTCAGTACCCGTCAGC 5160
Db GACTTGTGCTCGCTGTCTTGGAGGGTCTCTCTGAGTGATTCAGTACCCGTCAGC 5160
Qy GGGGGTCTTTCAATTTGG 5177
Db GGGGGTCTTTCAATTTGG 5177

RESULT 2

AAV81410

ID AAV81410 standard; DNA; 7699 BP.

XX

AC AAV81410;

XX

DT 17-OCT-2003 (revised)

XX

DT 26-APR-1999 (first entry)

XX

XX Plasmid pLAIN used in retroviral particle delivery.

XX

KW Plasmid pLAIN; retrovirus; gene transfer; mammary gland; milk; lactation;

KW

KW epithelial cell; trans-somatic animal; amylase; chicken; neo gene;

KW

KW cyclic; circular; ds.

XX

OS Moloney murine leukemia virus.

OS

OS Gallus sp.

OS

XX Chimeric.

XX

PN CA2224108-A.

XX

PD 05-SEP-1998.

XX

PF 24-FEB-1998; 98CA-02224108.

XX

PR 05-MAR-1997; 97CA-02199212.

XX

XX (MIAC) CANADA MIN AGRICULTURE.

PA

XX Gavora JS, Nguyen TH, Benkel BP, Falconer MM;

PI

XX WPI; 1999-035710/04.

DR

XX Production of a trans-somatic mammal for producing value-added milk - by

PT

PT transferring the gene into the secretory cells of the mammary gland using

PT

PT viral derived particles and their packaging cells.

XX

PS Example 1; Page 25-30; 53pp; English.

XX

This is the nucleotide sequence of pLAIN, a plasmid designed to carry the chicken amylase gene (as a marker protein) and various control sequences for use in retroviral particle delivery to bovine mammary epithelial cells. The plasmid has the amylase gene and a selectable marker neo gene both under control of the long terminal repeat promoter of the Moloney murine leukaemia virus. An internal ribosomal entry site (IRES) provides a translation initiation site within the transcript, allowing the downstream gene product to be produced. A method for producing a trans-somatic mammal for producing an altered composition of milk comprises:

CC

(1) providing a vector containing a DNA sequence encoding a valuable

CC

compound; (2) packaging the vector into a cell line to produce a trans-

CC

infecting particle; (3) preparing a solution comprising the trans-

CC

infecting particle; and (4) delivering the solution into the mammary

CC

gland to allow incorporation of the DNA into the secretory cells of the

CC

mammary gland. The new method is useful for producing milk with valuable

CC

compounds and/or a changed composition, to enhance its properties and/or

CC

marketability. The valuable compound is selected from: an antibody, an

CC

antibiotic, a blood clotting factor, galactosyltransferase, a growth

CC

factor, an oncoprotein, a hormone, a milk protein, a hormone receptor, a

CC

tumour suppressor protein, a vaccine, erythropoietin (claimed), a protein

CC

to aid in weight reduction, a protease inhibitor, an aging inhibitor and

CC

especially is a tissue plasminogen activator which may be extracted or

CC

the milk consumed directly as a therapeutic agent. The new method of

CC

producing trans-somatic animals has an increased trans-infection of

CC

mammary epithelial cells than prior art methods, resulting in an increase

CC

in the concentration and yield of the foreign protein in the milk. The

CC

new method requires only one infusion of the trans-infecting particle,

CC

and one or more genes can be trans-infected. (Updated on 17-OCT-2003 to

XX

standardise OS field)

XX

Sequence 7699 BP; 1814 A; 2023 C; 2014 G; 1848 T; 0 U; 0 Other;

| | | | | | |
|-----------------------|------|--|---------------|------------------|---------------------|
| Query Match | | 40.1%; | Score 2077.2; | DB 2; | Length 7699; |
| Best Local Similarity | | 64.3%; | Pred. No. 0; | | |
| Matches 3339; | | Conservative | 0; | Mismatches 1814; | Indels 38; Gaps 14; |
| Qy | 1 | TTTGAAGACCCACCCGCTAGTGGCAAGCTAGCTTAAGTAACGCCACTTTTGCAGGCAT | 60 | | |
| Db | 175 | TTTGAAGACCCACCCGCTAGTGGCAAGCTAGCTTAAGTAACGCCACTTTTGCAGGCAT | 234 | | |
| Qy | 61 | GGAAAAATACATAACTGAGAAATAGAAAAGTTTTCAGATCAAGGTCAGGAAACAAAGAAACAGC | 120 | | |
| Db | 235 | GGAAAAATACATAACTGAGAAATAGAAAAGTTTTCAGATCAAGGTCAGGAAACAAAGAAACAGC | 294 | | |
| Qy | 121 | TGAATACCAACAGGATATCTGTGGTAAGCGGTTCTGCCCCGGCTCAGGCCCAAGAACCA | 180 | | |
| Db | 295 | TGAATACCAACAGGATATCTGTGGTAAGCGGTTCTGCCCCGGCTCAGGCCCAAGAACCA | 354 | | |
| Qy | 181 | GATGAGACAGCTGAGTGTAGTGGCCCAACAGGATATCTGTGGTAAGCGGTTCTGCCCCGG | 240 | | |
| Db | 355 | GATGAGACAGCTGAGTGTAGTGGCCCAACAGGATATCTGTGGTAAGCGGTTCTGCCCCGG | 414 | | |
| Qy | 241 | CTCGGGGCCAAGAACAGATGTTCCCGAGATCGGTCAGACCCCTCAGCAGTTTCTAGTGAA | 300 | | |
| Db | 415 | CTCGGGGCCAAGAACAGATGTTCCCGAGATCGGTCAGACCCCTCAGCAGTTTCTAGTGAA | 474 | | |
| Qy | 301 | TCATCAGATGTTCCAGGTCGCCCAAGGACCTGAAAATGACCCCTGTACCTTATTTGAAC | 360 | | |
| Db | 475 | TCATCAGATGTTCCAGGTCGCCCAAGGACCTGAAAATGACCCCTGTACCTTATTTGAAC | 534 | | |
| Qy | 361 | TAAACAATCAGTTCGCTTCTCGCTTCTGTTCCGCGCTTCCGCTCTCCGAGCTCAATAAA | 420 | | |
| Db | 535 | TAAACAATCAGTTCGCTTCTCGCTTCTGTTCCGCGCTTCCGCTCTCCGAGCTCAATAAA | 594 | | |
| Qy | 421 | AGAGCCCAACACCCCTCACTCGGGCGGCAGTCTTCCGATAGACTGCGTCCCGCGGTAC | 480 | | |
| Db | 595 | AGAGCCCAACACCCCTCACTCGGGCGGCAGTCTTCCGATAGACTGCGTCCCGCGGTAC | 654 | | |
| Qy | 481 | CCGTATTCCCAATAAAGCCTCTGCTGTTTGCATCCGAATCGTGGTCTGCTGTTCTTG | 540 | | |
| Db | 655 | CCGTATTCCCAATAAAGCCTCTGCTGTTTGCATCCGAATCGTGGTCTGCTGTTCTTG | 714 | | |
| Qy | 541 | GGAGGGTCTCTCTGAGTGATGACTACCCAGACGGGGTCTTTCATTTGGGGCTCGT | 600 | | |
| Db | 715 | GGAGGGTCTCTCTGAGTGATGACTACCCAGACGGGGTCTTTCATTTGGGGCTCGT | 774 | | |
| Qy | 601 | CCGGGATTTGAGACCCCTGCGCCAGGACACCCAGCCACCCGGAGGTAAAGCTGCGC | 660 | | |
| Db | 775 | CCGGGATTTGAGACCCCTGCGCCAGGACACCCAGCCACCCGGAGGTAAAGCTGCGC | 834 | | |
| Qy | 661 | AGCAACTTATCTGTCTGTCGGATGTTCTAGTGTCTATGTTTGTATGTCGCTGCG | 720 | | |
| Db | 835 | AGCAACTTATCTGTCTGTCGGATGTTCTAGTGTCTATGTTTGTATGTCGCTGCG | 894 | | |
| Qy | 721 | TCTGTACTAGTTAGCTAACTAGCTCTGTATCTGCGGACCCGCTGGTGAACAGAGTT | 780 | | |
| Db | 895 | TCTGTACTAGTTAGCTAACTAGCTCTGTATCTGCGGACCCGCTGGTGAACAGAGTT | 954 | | |
| Qy | 781 | CTGAACACCCCGCGCAACCCCTGGGAGACGTCGCCAGGACCTTTGGGGCCGTTTGTGG | 840 | | |
| Db | 955 | CTGAACACCCCGCGCAACCCCTGGGAGACGTCGCCAGGACCTTTGGGGCCGTTTGTGG | 1014 | | |
| Qy | 841 | CCGACCTGAGGAAGGAGTGCATGTGGAAATCCGACCCCGTCAAGATATGTTGTTCTGTT | 900 | | |
| Db | 1015 | CCGACCTGAGGAAGGAGTGCATGTGGAAATCCGACCCCGTCAAGATATGTTGTTCTGTT | 1074 | | |
| Qy | 901 | AGGAGACGAGAACCTTAAACAGTTTCCCGCTCCGCTCTGAAATTTTCTTCGTTTGGAA | 960 | | |
| Db | 1075 | AGGAGACGAGAACCTTAAACAGTTTCCCGCTCCGCTCTGAAATTTTCTTCGTTTGGAA | 1134 | | |
| Qy | 961 | CCGAAGCGCGCTCTTGTCTGTGTCGACGCAAGCTTGGGCTGACAGTCTGATAGAGGA | 1020 | | |
| Db | 1135 | CCGAAGCGCGCTCTTGTCTGTGTCGACGCAAGCTTGGGCTGACAGTCTGATAGAGGA | 1194 | | |
| Qy | 1021 | TCAATTCGGCACGAGTAATCGTGCTGCTGCTTCTTAGGACATATGAAGTATGGCACAGT | 1080 | | |

| | | | | | |
|----|------|--|------|--|--|
| Db | 1195 | GACTGTGTTTCTGTGTAATTTGCTGAAATTAAGGCCAGACTGTTACCCTCCCTTAAGTTT | 1254 | | |
| Qy | 1081 | GGATGACTTTTCTGATCAGCAAGAGGACACTGACAGCTGTACAGAGTCTGTGAAGTTG | 1140 | | |
| Db | 1255 | GACCTTAGTCTACTTGGAAAGATGTCGAGGGATCGCTACACACCGTGGTAGATGTCAA | 1314 | | |
| Qy | 1141 | ATGCTCGCTCAGTGAAGCTTTGCTTCTCCCATCTTAAATAATGAGGAGGAGGAGGAGG | 1200 | | |
| Db | 1315 | GAAGAGAGCTTGGGTTACCTTCTGCTGCAAGATGCAACCTTTAAAGCTGGATGGCC | 1374 | | |
| Qy | 1201 | AGAGGATGAAGTCTTATAAAACCTGACATGATCAGCTTTATCTCTATGTTGTTGTTG | 1260 | | |
| Db | 1375 | GCGAGAGCGGACCTTTAAACCGAGACCTCATCACCAGGTTAAGATCAAGGCTCTT | 1431 | | |
| Qy | 1261 | TCGTGCCCATCATTTGGCATAGTGGCAGCTCAGCTCTGAAATGGAATGGAATTTGCA | 1320 | | |
| Db | 1432 | ACCTGGCCCGCATGGAACACCCAGACAGTCCCTCATCATCGTGACCTGGGAGGCTTGGC | 1491 | | |
| Qy | 1321 | CGGTTGGCTCAGTTTAATGTCAGATATATCTCAAGTCCGGAAGGCAAGAAATGGCAGTG | 1380 | | |
| Db | 1492 | TTTGAACCCCTCCCTGGGTCAAGCCCT-----TTGTACACCCCTAAGCCTCCGCTCCT | 1546 | | |
| Qy | 1381 | AAGATGAATGAGATTTTCGAGAGCTGTGATGGAAACGATGAGCAACATGGAAGAGCAAA | 1440 | | |
| Db | 1547 | CTTCTCCATCCGCCCCCTCTCTCCCTTTGAAACCTCTCGTTCGACCCCTCGATCC | 1606 | | |
| Qy | 1441 | TCAGATATCTTTAGATAATGAGCCCAATCTCTCTAGATGCTAAGAAATTTCCAAATTTCA | 1500 | | |
| Db | 1607 | TCCTTTTATCCAGCCCTCCTCTCTCTCTAGGCGCCGGAATTCGTTAACTCGACATGGA | 1666 | | |
| Qy | 1501 | GCATAACAACCTGATCAAGAAATTAATGATGTTCTTTTCCAGCTAAATTTCTTACTTCT | 1560 | | |
| Db | 1667 | GTCTTCTCTCTCTCGAGCTGTGGGCTTTGCTGGGACAGATACAAATCCCAACACTCAG | 1726 | | |
| Qy | 1561 | CCATCCAGG--AACATGAGAAATATCATAGGGATATCTCCAGTCTATTAGTAGTCTGAA | 1618 | | |
| Db | 1727 | GCTGGGAGGACATCTATCTGTCATCTCTTTGAATGGGCTGGGCGGACATTTGCATGGAG | 1786 | | |
| Qy | 1619 | CACCACAGTACTTGAATTTGAGTTTCACTATTTGAACACTGAAATGGCAGAGTCCCAAGAA | 1678 | | |
| Db | 1787 | TGCGAACACTATTAGTCTCTTAATGGGTTTGGAGAGTTTCAAGTTTCTCTCCAAATGAA | 1846 | | |
| Qy | 1679 | TGCAATTTAAACAAACAGAGGAGATGCGTAAATTAAGAGGAGCGTATATCAATGATCAGC | 1738 | | |
| Db | 1847 | AACATTTGCTACTTAATCCGAACAGGCCCTCGTGGGGAAGATACCAGCCCTCAGCTAC | 1906 | | |
| Qy | 1739 | AGAAATTAAGTCTCTAGATGAAACAAAGTATATTTGGAAACAGGAAATAAAGGGGAAAT | 1798 | | |
| Db | 1907 | A-AGATCTGCACCTCGATCGGCCAATGAAATGAATTCAGAGACATGGTGACAGATGCA | 1965 | | |
| Qy | 1799 | GAACTGTTGNAATTAATCACTAATGATCTGAGGCTGAAGGATTTGGGNAACATTTCTCAGAC | 1858 | | |
| Db | 1966 | CAATGTTGGAGTTGCTATTTATGTTGATGCTGTTGTTCAATCAGATGTGTGATCTATGGG | 2025 | | |
| Qy | 1859 | ATTGAAAAATATCACTTTACTCCAAAGGTGCGAGAAAGTGTCTGCTGACTGGGAAATGAC | 1918 | | |
| Db | 2026 | TGGCAGCGGACCCACTCAACATGTTGGAGCTATTTCAACACCGGACTAGAGATTTTCC | 2085 | | |
| Qy | 1919 | CAACGATCTGGGCTCAACATGACCATCTGGGGCTGTGAAACAGAGAGGTGAATTTCAAGG | 1978 | | |
| Db | 2086 | CGCTGTCCGCTACTCTGCTGGGATTTCAATGACGGCAATGTCACTGCAAGTGGAGA | 2145 | | |
| Qy | 1979 | CACCTACATCACAGCCCTTAACGCCACATCAATGAGATCAAGAGTCAACAGTCCACATGG | 2038 | | |
| Db | 2146 | CATCGAAAAATTAATGGGACATGATATCAGGTCGGGATTTGCAAGTTGTCAGCTCTTGA | 2205 | | |
| Qy | 2039 | GACAAAAACACCATCAACAGAGGACCCAGGCCACCTTTGGGCTTCA-CGCTCAATTTGA | 2097 | | |
| Db | 2206 | TCGCTCTCTGGAGAGGACTATGTACGCTCAACAAATTTGACGCTACATGATCACTCAT | 2265 | | |
| Qy | 2098 | AGTTTTCAGAGTCCACCACTGTCTTTCAGCGGCCAGTGTCTTATAGACAGGAATGGGAGG | 2157 | | |

Db 2266 TGATATGGGTAGCAGGGTTCCGGATCGATGCTGCCAAGCATATGTGGCCAGGGGACAT 2325
Qy 2158 AGGTCTGAAGACCATGTGGCTGTGGCTCAAGTGTAAATGACATTTGGTGTATGACTGGA 2217
Db 2326 AAGAGCGTTTCTGACAAACCTGCACGATCTAAATACTCAGTGGTTTTCTAGCAGGAACGAA 2385
Qy 2218 AAGCTACAGGGTC - GGCAATCAACATCTTCACTCGCTCGCCACACAGAAAGAGTGTAGTG 2276
Db 2386 ACCCTTTATTACCAAGAGGTAAATTGACTTGGGAGGAGAGCCAAATCA CAGGCACTCAGTA 2445
Qy 2277 AGTGACCAAGTCTCTCTG --- GACTCCAGGTGAAAAGGAGATAGAGGCCCTCTCTGAC 2333
Db 2446 CTTTGGGAATGGCCGCGTGACAGAAATCAAGATATGGTGCCTTAACTTGGGGACGGTGATCCG 2505
Qy 2334 AAAATGGTATACAGGCTTTCAGGTCTAAATAGGTACTCCAGGTCTTAAAGGTGATCGG 2393
Db 2506 GAAGTGGNATCGAGAGAAGATGGCTACTTAAAGAACTTGGGGAAGGCTGGGCTTTGT 2565
Qy 2394 GGGATCTCTGTTTAACTTGGAGTTTCGAGATTCCTCAGGAC - -CAATGGGGAAGACCCGGA 2451
Db 2566 GCCTTCTGACAGAGCCCTGGTGTGTGGATAAACACAGACAAACAGCGGGGCGACGGGC 2625
Qy 2452 AGCCAGGACTTAATGACAAAAGCCAGAAAGGAGAAAAGGAGTGGAAAGCATGCAAA 2511
Db 2626 AGCGGAGCTTCCATTTCTTACTTTCTGGGATGCCAGGCTTTATAAAATGGCGGTTGGTTT 2685
Qy 2512 GACAATCTAATACAGTCCGACTGTGGTGGTGCGAGCGCCCTCACGAAGGCAGATGGAGA 2571
Db 2686 CATGCTCCTCATCCGTACGGGTTCA CACGGGTGATGTCAAGTTATCGTTGGCCAAAGATA 2745
Qy 2572 TTTTTCAGAAAGCCAGTGGGTACGGTGTGTGACGACCGCTGGGAACTCGTGGAGGAC 2631
Db 2746 TTTTCGAAAACCGAGTGGATGTTAACGACTGGGTGGGACCACCAAGTAAC TCGGACGGATC 2805
Qy 2632 T-----GGTGTCTGACAGAGCTTGGGATACAAAGGTGTTCAAAGTGTGCATATAAGCGACT 2687
Db 2806 GACGAAGTCCGTTACAAATCAACGCAGACACTACTGTGGCAATGACTGGGTCTCGGAACA 2865
Qy 2688 TATTTTGGAAAAGGTACGGGTCCTAATATGGCTGAATGAATGATTTTGTTCGGGAAAGAG 2747
Db 2866 TCGTGGCGGCAAAATAGG-----ACATGGTTATCTTCCGTAATGTGTAGACGGTCAGC 2921
Qy 2748 TCATCCATTTGAAGAGTGCAGAAATTAGACAGTGGGGTGTGAGAGCTGTTCGCACAGCAAA 2807
Db 2922 CTTTCTCAAACTGGTGGGACAAACGGAGCAATCAAGTAGCTTTTCGGTCGGGCGACAGAG 2981
Qy 2808 GATGTGGGGGTCACTTGTGACCTACATAATGCAATCATATTTTCATTCACATTTTTTAA 2867
Db 2982 GCTTCATTTGCTTTAATAATGATGACTGGTATATGAATGTGCAATTTGCAAACTGGTCTGC 3041
Qy 2868 CTGTTATAAGTGAATTTTTTCTTCTTCTTCACTTAAATCAGCTTAATTAATAATTTAAGA 2927
Db 3042 CTGCTGGAACTTACTCGCATGTTATTTCTGACAAAAGGAAGGAGCGGTGCTGTACTGGAA 3101
Qy 2928 AACTAAGAAATTTTATCCACAGAAAAGGAATATTTAAATAATCACTGGATAACATATAAAA 2987
Db 3102 AGCAGGTGTACGTTCTCGATGGAAAGGCCAATTTCCAGATTAGTAAACAGCATGAG 3161
Qy 2988 TAGCTTCATATTTGCTTCAATAACAGAAACCATTTCACTTCTCTAGGTTTTTAAAGTGGC 3047
Db 3162 ATCCATTTGTTGCAATTCACGTTGATGCCAAGTTATAAGCTTTCGAGGATCCACTAGTAAC 3221
Qy 3048 TCGTGGCAATTTGATCCCTCAGGATATAGTATGTTTCGCTTTTTCATAGGGGGGGGAAA 3107
Db 3222 GGCGCCAGTGTGCTGGAAATTCGG-----CTTGTGCAATCTAGGGCGGCCAATTCGCC 3276
Qy 3108 TGTAGTCTTATGCAATACTCTTGTAGTCTTGCAACATGGTAACGATGAGTTAGCAACATG 3167
Db 3277 CCTCTCCCCCCCCCTTAAGCTTACTGGCGAAGCCGCTTGGAAATAGGCCGCTGTGTG 3336
Qy 3168 CCTTACAGGAGAGAAAAGCACCCTGCATGCCGATTTGGTGGAAAGTAAAGTGTGTACGATC 3227
Db 3337 TTTGTCTATATGTGAATTTTCCACCATATTTGCCGTCTTTTGGCAA-----TGTGAGGGCC 3390

Qy 3228 GTGCTTATTAGGAAGCAACAGACGGGTCTGACATGGAATTGGACGAACCACTGAATTCC 3287
Db 3391 CGGAAACCTGGCCCTGTCTTTCTTACGAGCATTTCTAGGGGTCTTTCCCTCTCGCCAAA 3450
Qy 3288 GCATTTGACAGATATTTAATAGTGCCTAGTCTGATACAGAAACGCAATTTGACCAT 3347
Db 3451 GGAATGCAAGGTCTGTTGAATGCTGTAAGGAAGCAGTCTCTGGAAGCTTCTTGAAGA 3510
Qy 3348 TCACCAATTTGGTGTGCACCTCCAGCTTACCGTCCGCAAGCACTCAGCGGCGCAAGGG 3407
Db 3511 CAAACAAAGCTGTGTAGCGACCTTTTGACGAGCGGAAACCCCACTCTGGCGACAGGTGC 3570
Qy 3408 CTGCTAAAGGAAGCGGAAACAGCTAGAAAGCGAGTCCGACAGAAACGCTGCTACCCCGCAT 3467
Db 3571 CTCTGGGCCCAAAAGCCACGCTGTATAGATACACTGCAAAAGCGGCAACACCCAGTGC 3630
Qy 3468 GAATGTCACTATCTGGCTATCTGGACA - AGGGAAAACGCAAGCGCAAGAAAGCAAGG 3526
Db 3631 CACGTTGTGAGTTGGA TAGTTGTGGAAGAGTCAAATGGCTCTCTCAAGCGTAGTCAAC 3690
Qy 3527 TAGCTTGCAGTGGGCTTACATGGGATAGCTAGACTGGGCGGTTTTATGGACAGCAAGCG 3586
Db 3691 AAGGGGCTGAAGGATGCCAGAAAGGTACCCCATTTGTATGGGAATCTGATCTGGGGCCTCG 3750
Qy 3587 AACCGGAATTTGCAGCTGGGCGGCCCTCTGCTAAGGTTGGGAAGCCCTGCAAAAGTAAACT 3646
Db 3751 GTGCACATGCTTTACATGTGTTTAGTCGAGGTTAAAGAACTCTAGGCCCCCGCAACCAC 3810
Qy 3647 GGAATGGCTTTTTCGCCCAAGGATCTGATGGCGCAGGGATCAAGATCTGATCAAGAGA 3706
Db 3811 GGGAGCTGGTTTTTCTTTGAAAACACAGATGATAAGCTTGGCCACAAACC CGGGATAATTC 3870
Qy 3707 CAGGATGAGGATCGTTTCGCGATGATTTGAAACAAGATGATTTGCAAGCAATCGGCTCTCATG 3766
Db 3871 CTGCAGCCAAATATGGGATCGGCCAATTTGAACAAGATGATTTGCAAGGTTTCTCCGGCCG 3930
Qy 3767 CTTGGGTGGAGAGGCTATTCGGCTATGACTGGGCAACACAGACAATCGGCTGCTCATG 3826
Db 3931 CTTGGGTGGAGAGGCTATTCGGCTATGACTGGGCAACACAGACATCGGCTGCTCATG 3990
Qy 3827 CCGCCGTTTCCGGCTGTACGCGAGGGCGCCGGTTCTTTTGTCAAGCAAGCACTGT 3886
Db 3991 CCGCGGTTTCCGGCTGTACGCGCAGGGCGCCGGTTCTTTTGTCAAGACCCGACCTGT 4050
Qy 3887 CCGGTGCCCTGAATGAATGAACTGCAAGCAGGAGCGCGGCTATCGTGGCTGCCACGACGG 3946
Db 4051 CCGGTGCCCTGAATGAATGAACTGCAAGCAGGAGCGCGGCTATCGTGGCTGCCACGACGG 4110
Qy 3947 GCGTTCTTCCGACGCTGTCTCGACGTTGTCTGAAAGCGGGAAGGAGCTTGGCTGCTAT 4006
Db 4111 GCGTTCTTCCGACGCTGTCTCGACGTTGTCTGAAAGCGGGAAGGAGCTTGGCTGCTAT 4170
Qy 4007 TGGGCGAAGTCCCGGGCAGGATCTCTGTCTCATCTCACTTGTCTCTGCCGAGAAAGTAT 4066
Db 4171 TGGGCGAAGTCCCGGGCAGGATCTCTGTCTCATCTCACTTGTCTCTGCCGAGAAAGTAT 4230
Qy 4067 CCATCATGGCTGATGCAATCGCGGCGCTGATACGCTTGTATCCGGCTACCTGCCCATTCG 4126
Db 4231 CCATCATGGCTGATGCAATCGCGGCGCTGATACGCTTGTATCCGGCTACCTGCCCATTCG 4290
Qy 4127 ACCACCAAGGAAACATCTGCATCGAGCGAGCAGCTACTCCGATGGAAGCGCGGTCTTGTGC 4186
Db 4291 ACCACCAAGGAAACATCTGCATCGAGCGAGCAGCTACTCCGATGGAAGCGCGGTCTTGTGC 4350
Qy 4187 ATCAGGATGATCTGGACGAAGAGCATCAGGGGCTTCGGCGCAGCCGCAACTGTTCTGCCAGGC 4246
Db 4351 ATCAGGATGATCTGGACGAAGAGCATCAGGGGCTTCGGCGCAGCCGCAACTGTTCTGCCAGGC 4410
Qy 4247 TCAAGGGCGCATGCCCGACGGCGAGATCTGCTGCTGACCCATGGCGATGCCCTGCTTGC 4306
Db 4411 TCAAGGGCGCATGCCCGACGGCGAGTATCTGCTGTTGACCCATGGCGATGCCCTGCTTGC 4470

QY 4307 CGAATATCATGGTGGAAATGCGCGCTTTCTGGATTGATCGACTGTGGCGGCTGGGTG 4366
Db 4471 CGAATATCATGGTGGAAATGCGCGCTTTCTGGATTGATCGACTGTGGCGGCTGGGTG 4530
QY 4367 TGGGGGACCGCTATCAGGACATAGCGTTGGCTACCGGTGATATTGCTGAAGAGCTTGGCG 4426
Db 4531 TGGGGGACCGCTATCAGGACATAGCGTTGGCTACCGGTGATATTGCTGAAGAGCTTGGCG 4590
QY 4427 GCGAATGGCTGACCGCTTCTCGTGTCTTACGGTATCGCGCTCCCGATTTCGAGGCGCA 4486
Db 4591 GCGAATGGCTGACCGCTTCTCGTGTCTTACGGTATCGCGCTCCCGATTTCGAGGCGCA 4650
QY 4487 TCGCTCTCTATCGGCTTCTTACGAGTCTTCTCGAGCGGACTCTGGGTTTCGATAAAAT 4546
Db 4651 TCGCTCTCTATCGGCTTCTTACGAGTCTTCTCGAGCGGACTCTGGGTTTCGATAAAAT 4710
QY 4547 AAAAGATTTTATTAGTCTCCAGAAAAGGGGGGGAATGAAAGACCCACCTGTAGTTTG 4606
Db 4711 AAAAGATTTTATTAGTCTCCAGAAAAGGGGGGGAATGAAAGACCCACCTGTAGTTTG 4770
QY 4607 GCAAGCTAGCTTAAAGTAAAGCCATTTTGCAGGATGGAATAAATACATAAAGTGAATAG 4666
Db 4771 GCAAGCTAGCTTAAAGTAAAGCCATTTTGCAGGATGGAATAAATACATAAAGTGAATAG 4830
QY 4667 AGAAGTTTCAGATCAAGGTGAGAACAGATGGAACAGCTGAATATGGCCAAACAGGATAT 4726
Db 4831 AGAAGTTTCAGATCAAGGTGAGAACAGATGGAACAGCTGAATATGGCCAAACAGGATAT 4890
QY 4727 CTGTGGTAAGCAGTTCTGCGCCCGCTCAGGCGCAAGAACAGATGGAACAGCTGAATATG 4786
Db 4891 CTGTGGTAAGCAGTTCTGCGCCCGCTCAGGCGCAAGAACAGATGGAACAGCTGAATATG 4950
QY 4787 GCGCAAAACAGGATATCTGTGTAAAGCAGTTCTGCGCCCGCTCAGGCGCAAGAACAGATG 4846
Db 4951 GCGCAAAACAGGATATCTGTGTAAAGCAGTTCTGCGCCCGCTCAGGCGCAAGAACAGATG 5010
QY 4847 GTCCCCAGATCGGTCAGCCCTCAGCAGTTCTTAGAAGACCATCAGATGTTTCCAGGCT 4906
Db 5011 GTCCCCAGATCGGTCAGCCCTCAGCAGTTCTTAGAAGACCATCAGATGTTTCCAGGCT 5070
QY 4907 GCGCCAGAGGACCTGAATAGCCCTGTGCTTATTGAACTAACCAATCAGTTGCTCTCTC 4966
Db 5071 GCGCCAGAGGACCTGAATAGCCCTGTGCTTATTGAACTAACCAATCAGTTGCTCTCTC 5130
QY 4967 GCTTCTGTTCGCGGCTTCTGCTCCCGAGCTCAATAAAGAGCGCCCAACCCCTCACTC 5026
Db 5131 GCTTCTGTTCGCGGCTTCTGCTCCCGAGCTCAATAAAGAGCGCCCAACCCCTCACTC 5190
QY 5027 GGGGCGCCAGTCTCCGATTCAGTTCGAGTCGCGCGGTTACCGGTGATTCCAATAAACCCCTC 5086
Db 5191 GGGGCGCCAGTCTCCGATTCAGTTCGAGTCGCGCGGTTACCGGTGATTCCAATAAACCCCTC 5250
QY 5087 TTGCAGTTGCATCCGACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5146
Db 5251 TTGCAGTTGCATCCGACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5310
QY 5147 GACTACCCGTCAGCGGGGCTCTTTCATTTGG 5177
Db 5311 GACTACCCGTCAGCGGGGCTCTTTCATTTGG 5341

RESULT 3

ADQ80675

ID ADQ80675 standard; DNA; 7436 BP.

XX AC ADQ80675;

XX DT 21-OCT-2004 (first entry)

XX DE Bcl-XL expression directing retroviral vector, pPB-bcl-XL-Sn.

XX KW Survival; neuron; tyrosine hydroxylase; tyrosine 3-monooxygenase; TH; anti-apoptotic; Bcl-XL; neurological disorder; neuroprotective;

nootropic; antiparkinsonian; transplantation; drug screening; gene profiling; CNS disorder; neurodegenerative disease; pPB-bcl-XL-Sn; ds.

Unidentified.

W02004062554-A2.

29-JUL-2004.

07-JAN-2004; 2004WO-DK000008.

08-JAN-2003; 2003US-0438719P.

11-APR-2003; 2003DK-0000581.

22-APR-2003; 2003US-0464546P.

(NSGE-) NSGENE AS.

Martinez-Serrano A, Liste I, Villa A;

WPI; 2004-544027/52.

Enhancing the survival of neurons or cells expressing tyrosine hydroxylase (TH) for treating neurodegenerative disorders, comprises contacting neurons or TH expressing cells with Bcl-XL or its functional equivalent.

Disclosure; SEQ ID NO 6; 108pp; English.

The invention relates to a novel method for enhancing the survival of neurons and/or of cells expressing tyrosine hydroxylase (EC 1.14.16.2 - Tyrosine 3-monooxygenase) (TH +). The method comprises contacting a population of cells with Bcl-XL or its functional equivalent, where the population of cells is selected from: neurons or cells capable of differentiating into neurons; or TH expressing cells or cells capable of differentiating into TH expressing cells. The invention further comprises: a composition of cells obtainable by the method above; a composition of isolated mammalian cells overexpressing the anti-apoptotic Bcl-XL protein; a neural progenitor cell; a differentiated dopaminergic neuron; an implantable cell culture device comprising: a semi-permeable membrane permitting the diffusion of a biologically active protein through it; and a composition of cells selected from above; a lentiviral vector particle being produced based on a lentiviral transfer vector; enhancing the survival of TH + cells in vivo; a retroviral particle being produced based on a retroviral transfer vector; enhancing the survival of in vivo differentiated dopaminergic neurons; a packaging cell line capable of producing an infective vector particle; a packaging cell line capable of producing an infective vector particle; treatment of a neurological disorder; a fusion protein comprising the Bcl-XL sequence comprising 233 amino acids ADQ80670 or its functional equivalent and a membrane translocation signal; an expression vector comprising a polynucleotide sequence coding for the fusion protein and a promoter sequence capable of directing the expression of the fusion protein in a host cell; a host cell comprising the expression vector; and producing the fusion protein. The compositions of the invention have neuroprotective, neurotropic, and antiparkinsonian activities. The cells are useful for transplantation, drug screening, gene profiling, or for the preparation of a medication useful for the treatment of a CNS disorder. The CNS disorder is a neurodegenerative disease involving lesioned and traumatic neurons, including traumatic lesions of peripheral nerves, the medulla, the spinal cord, cerebral ischaemic neuronal damage, neuropathy, peripheral neuropathy, Alzheimer's disease, Huntington's disease, Parkinson's disease, amyotrophic lateral sclerosis, or memory impairment connected to dementia. The method is useful for enhancing the survival of neurons and/or of cells expressing tyrosine hydroxylase for the treatment of neurodegenerative disorders. This sequence represents an expression construct for directing the expression of Bcl-XL in transduced cells, used in the method for increasing the survival rate of neurons of the invention.

Sequence 7436 BP; 1744 A; 1993 C; 1939 G; 1760 T; 0 U; 0 Other;

Query Match 38.1%; Score 1972.6; DB 13; Length 7436;

Best Local Similarity 65.1%; Pred. No. 0;
Matches 3371; Conservative 0; Mismatches 1529; Indels 281; Gaps 20;

| | | | |
|----|------|---|------|
| Qy | 1 | TTTGAAGACCCACCCGTAGGTGGCAGCTAGCTTAAGTAAACGCCACTTTGCAAGGCAT | 60 |
| Db | 1 | TTTGAAGACCCACCCGTAGGTGGCAGCTAGCTTAAGTAAACGCCACTTTGCAAGGCAT | 60 |
| Qy | 61 | GGAAAAATACATACTAGAGATAGAAAAGTTTCAGATCAAGGTTCAGGACAAAGAAACAGC | 120 |
| Db | 61 | GGAAAAATACATACTAGAGATAGAAAAGTTTCAGATCAAGGTTCAGGACAAAGAAACAGC | 120 |
| Qy | 121 | TGAATACCAACACAGGATATCTGTGGTAAAGCGTTCTCTGCCCGGCTCAGGGCCCAAGAAC | 180 |
| Db | 121 | TGAATACCAACACAGGATATCTGTGGTAAAGCGTTCTCTGCCCGGCTCAGGGCCCAAGAAC | 180 |
| Qy | 181 | GATGAGACAGTGTGATGGGCGGCAACACAGATATCTGTGTAGAGCAGTTCTCTGCCCGG | 240 |
| Db | 181 | GATGAGACAGTGTGATGGGCGGCAACACAGATATCTGTGTAGAGCAGTTCTCTGCCCGG | 240 |
| Qy | 241 | CTCGGGCCAAAGAACAGATGGTCCAGATGCGGTCCAGCCCTCAGCAGTTTCTAGTGAA | 300 |
| Db | 241 | CTCGGGCCAAAGAACAGATGGTCCAGATGCGGTCCAGCCCTCAGCAGTTTCTAGTGAA | 300 |
| Qy | 301 | TCATCAGATGTTCCAGGGTCCCAAGACCTGAAATGACCTGTACCTTTATTTGAAC | 360 |
| Db | 301 | TCATCAGATGTTCCAGGGTCCCAAGACCTGAAATGACCTGTACCTTTATTTGAAC | 360 |
| Qy | 361 | TAAACCATCAGTTCGCTCTCTGCTTCTGTTCCGCGGCTTCCGCTCCGAGCTCAATAA | 420 |
| Db | 361 | TAAACCATCAGTTCGCTCTCTGCTTCTGTTCCGCGGCTTCCGCTCCGAGCTCAATAA | 420 |
| Qy | 421 | AGAGCCCAACACCTCACTCGGGCGGCGAGTCTTCGATAGACTGCGTCCCGGGTAC | 480 |
| Db | 421 | AGAGCCCAACACCTCACTCGGGCGGCGAGTCTTCGATAGACTGCGTCCCGGGTAC | 480 |
| Qy | 481 | CCGTAATCCCAATAAAGCCTTCTGCTGTTGCATCCGAATCGTGTCTCGCTGTTCTCTG | 540 |
| Db | 481 | CCGTAATCCCAATAAAGCCTTCTGCTGTTGCATCCGAATCGTGTCTCGCTGTTCTCTG | 540 |
| Qy | 541 | GGAGGCTCTCTCTGAGTGTAGTACTACCCAGCAGGGGGTCTTTCATTTGGGGGCTCGT | 600 |
| Db | 541 | GGAGGCTCTCTCTGAGTGTAGTACTACCCAGCAGGGGGTCTTTCATTTGGGGGCTCGT | 600 |
| Qy | 601 | CCGGATTTGAGACCCCTGCGCAGGACACCCAGACCCACCCAGGGAGTAACTGCGC | 660 |
| Db | 601 | CCGGATTTGAGACCCCTGCGCAGGACACCCAGACCCACCCAGGGAGTAACTGCGC | 660 |
| Qy | 661 | AGCAACTTATCTGTCTGTCTGCGATGTTGTCTAGTGTCTATGTTGATGTTATGCGCCTGCG | 720 |
| Db | 661 | AGCAACTTATCTGTCTGTCTGCGATGTTGTCTAGTGTCTATGTTGATGTTATGCGCCTGCG | 720 |
| Qy | 721 | TCTGTACTAGTTAGCTAACTAGCTCTGTATCTGTGGGACCCGTTGGTGAACCTGACGAGTT | 780 |
| Db | 721 | TCTGTACTAGTTAGCTAACTAGCTCTGTATCTGTGGGACCCGTTGGTGAACCTGACGAGTT | 780 |
| Qy | 781 | CTGAACACCCCGGCGCAACCTCGGAGACGTTCCAGGACCTTTGGGGCGGTTTTTGTGG | 840 |
| Db | 781 | CTGAACACCCCGGCGCAACCTCGGAGACGTTCCAGGACCTTTGGGGCGGTTTTTGTGG | 840 |
| Qy | 841 | CCGACCTGTAGGAAGGAGTGTGGAATCCGACCCCGTCAGGATATGTGGTTCTGGT | 900 |
| Db | 841 | CCGACCTGTAGGAAGGAGTGTGGAATCCGACCCCGTCAGGATATGTGGTTCTGGT | 900 |
| Qy | 901 | AGGAGACAGAACCTAAACAGTTTCCCGCTCCGCTCTGGAATTTTTCGTTTCGGTTGGAA | 960 |
| Db | 901 | AGGAGACAGAACCTAAACAGTTTCCCGCTCCGCTCTGGAATTTTTCGTTTCGGTTGGAA | 960 |
| Qy | 961 | CCGAAGCCGCGCTCTGTCTGTGTCAGCCCAAGCTTGGGCTGCAGCTCGACTCTAGAGGA | 1020 |
| Db | 961 | CCGAAGCCGCGCTCTGTCTGTGTCAGCCCAAGCTTGGGCTGCAGCTCGACTCTAGAGGA | 988 |
| Qy | 1021 | TCAATTCGGCAGTAAATCCGCTGCTCCGCTCTTTAGGACATATGAAGTATGGCAGCT | 1080 |
| Db | | | |

| | | | |
|----|------|---|------|
| Db | 989 | -----CGCTGCAGCATCGTTTCTGTGTGTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT | 1028 |
| Qy | 1081 | GGGATGACTTCTCTGATCAGCAGAGGACACTGACAGCTGTACAGAGTCTCTGAGGTTTCG | 1140 |
| Db | 1029 | TTCTGTATTTGTCTGAAAAATTAGGGCCAGACTGTTTACC-----ACTCCCTTAAGTTTG | 1081 |
| Qy | 1141 | ATGCTCCGCTCAGTGACAGCTTTGCTTCTCCCATCTCTCAAAATGCGGCCAACTCTTCAAG | 1200 |
| Db | 1082 | ACCTTAGTCTACTTGGAAAGATGTGAGGGGATCGCTCACAACAGTCCGTAGATGTCAAG | 1141 |
| Qy | 1201 | AGAGATGAAGTCTTATAAACTGACACTGATCACTTTTATCTCATTTGTGTTGTAGTTC | 1260 |
| Db | 1142 | AAGAGAGCTTTGGGTTA-CCTTCTGCTCTGCAGATGCGCAACCTTTAAACGTCGAGTGCC | 1200 |
| Qy | 1261 | TCGTGCCCATCATTTGGCATATGTGCAGCTCAGCTCTCTGAAATGGAACAGGAATTCGA | 1320 |
| Db | 1201 | GCAGAGCGGACCTTTTAAACCGAGACCTCATCACCCAGGTTAAGATCAAGGTCCTTTTACC | 1260 |
| Qy | 1321 | CGGTTGGCTCAGTTAATGTCAGATATATCTCAAGTCCGGAAGGCAAGGAATGCGCAGTG | 1380 |
| Db | 1261 | TGGCCCGCATGGACACCCAGACAGGTCCTCATCATCTGTGACCT--GGGAAGCCTTGGC | 1317 |
| Qy | 1381 | AAGATGAATAGAGATTTTCGAGAAAGCTGTGATGGAAACGATGAGCAACATGGAAGCAGAA | 1440 |
| Db | 1318 | TTTTGACCCCTCCCTGGGTCAAGCCTTTGTACACCCCTAAGCCTCGCCTCTCTCTTCC | 1377 |
| Qy | 1441 | TCAGATATCTTTCAGATAATGAAGCCAACTCTCTAGATGCTAAGAAATTTCCAAATTTCA | 1500 |
| Db | 1378 | TCCATCGCCCGGCTCTCTCCCTTGAACCTCTCTGTTTCGACCCCGCTCGATCTCCCT | 1437 |
| Qy | 1501 | GCATAACCACTGATCAAGATTTAATGATGTTCTTTTCCAGCTAAATTTCTTACTTCTCT | 1560 |
| Db | 1438 | TTATCCAGCCTCACTCTCTCTTAGGCGCGGAATTCGATGTACGGGCCAGATATAGGC | 1497 |
| Qy | 1561 | CCATCCAGGAACATGAGAAATATCATAGGGGATCTCCAAAGTCAATTAGTCTGAGACA | 1620 |
| Db | 1498 | GTTGACATTGATTTAGTCTAGTTAATTAAGTAATCAATTTACGGGGTCAATTAGTTCATA | 1557 |
| Qy | 1621 | CCAGTACTTGTATTTGAGTTCAGTATTGAACACTGAAATGGCAGAGTCCAAGAGAATG | 1680 |
| Db | 1558 | GCCATATATGAGGTTCCGGTTACATACTTACGGTAAATGGCCCGCTGGCTGACCGC | 1617 |
| Qy | 1681 | CATTAAACCAAGAGAGAGATGCTAAATTTAGAGGAGCGTATATACAAATGCATCAGCAG | 1740 |
| Db | 1618 | CCAAGACCCCGCCCATTCAGCTCAATAATGACGTATGTTCCCATAGTAACGCCAATAG | 1677 |
| Qy | 1741 | AAATTAAGTCTAGATGAAAAACAGTATATTTGGAAACAGGAATAAAGGGGAAATGA | 1800 |
| Db | 1678 | GGACTTTCCATTCAGCTCAATGGGTGACTATTACGGTAAACTGCCCCACTTGGCAGTAC | 1737 |
| Qy | 1801 | AACGTTGAATATATCATTAATGATCTGAGGCTGAAGGATTTGGGAACATTTCTCAGACAT | 1860 |
| Db | 1738 | ATCAAGTGTATCATATGCCAAGTACGCCCTTATTTGACGTCAATGACGGTAAATGGGCCG | 1797 |
| Qy | 1861 | TGAAAAATATACCTTTACTCCAAAGTCCAGAAAGTCTCGCTGACTGGGAAATGGACCA | 1920 |
| Db | 1798 | CCTGGCAT---TATGCCCCAGTACATGACCTTATGGGACTTTCTTACTTTGGCAGTACATCT | 1854 |
| Qy | 1921 | ACGATCTGGGCTCAACATGACCATCGGGGCTGTGAACAGCAGAGGTGAATTCACAGGCA | 1980 |
| Db | 1855 | ACGTTATAGTTCATCGCTATTACCATGGTG-----ATGCGGTTTTTGGCAGTAC | 1901 |
| Qy | 1981 | CCTACATCAGCGGTAAACAGCCACATCAAAATGAGATCAAGAGTACCACCTGATCGGGA | 2040 |
| Db | 1902 | ATCAATGGGCGTGGATAGCGGTTTGACTCAGGGGATTTTCCAGGTCTCCACCCCATTTGAC | 1961 |
| Qy | 2041 | CACAAAACACCATCAACAGAGGACCCAGCCCACTTTGGCTTCCCGTCAATTTGGAGT | 2100 |
| Db | 1962 | GTCAATGGGAGTT-----TGTTTTGGCACCAAAAATCAACGGGACTT | 2002 |
| Qy | 2101 | TTTCAGAGTCCACACTGTCTTTCAGGGCCAGTGTCTTTATAGACAGGAATGGGAAGGAGG | 2160 |
| Db | 2003 | TCCAAAATGTCGTAAC---AACTCCGCCCCCATTTGACGCAAAATGGGCGGTAGGCGTGTACG | 2059 |

QY 2161 TCCTGAAGCCATGTGGCTGCTGGCTCAAGTGTAAATGACATTTGGTGTATGACTGGAAG 2220
Db |||||
QY 2060 GTGGAGGTCTATAT--AAGCAGAGCTCTGGCTAACTAGAGAACCCACTGCTTACTGG 2117
Db |||||
QY 2221 CTACACAGGTGGCATCAACATCTTCACTCGCTGCGCACACAGAGAGTGTAGTGTG 2280
Db |||||
QY 2118 CTTATCGAANTTAATACGAC---TCACTATAGGAGACCCCAAGCTGGCTAGCGTTTAA 2173
Db |||||
QY 2281 ACCAAGTCTCTCTGGAATCCAGGTGMAAAAGGAGATAGAGGCCCTCTGACACAAATGG 2340
Db |||||
QY 2174 CTTAAGCTTGGTACCGAGCTCGATCCACTAGTCCAGTGTGTGGAATTTCTGAGATATC 2233
Db |||||
QY 2341 TATACAGGCTTTCCAGGTCTAATAGTATCTCCAGGTCTTAAAGGTGATCGGGGGATCT 2400
Db |||||
QY 2234 CAGCACAGTGGCGCGCTCGATATAAATGTCTCAGAGCAACCGGAGCTGGTGGTGCAC 2293
Db |||||
QY 2401 CTGGTTTACCTGGAGTTCCAGGATTCGCCAGACCAATGSGG--AAGACCGGGAAGCCAGG 2458
Db |||||
QY 2294 TTTCTCTCTACAGCTCTCCAGAAAGGATACAGTGGAGTCAGTTTAGCGATGTGAA 2353
Db |||||
QY 2459 ACTTAATGGACAAAAGGCCAGAGAGGAGAGAAAAGGAGTGGAAAGCATGCAAGACAAATC 2518
Db |||||
QY 2354 GAGAACAGGACTGAAGCCCCAGAGAACTGAACAGAAAGGAGACCCCACTGCGCATC 2413
Db |||||
QY 2519 TAATACAGTCCGACTGGTGGTGGAGCGGCCCTCAAGAAAGGAGAGTGGAGATTTTCA 2578
Db |||||
QY 2414 AATGGCAACCCATCTGACCTCGCGATAGCCCGCGTGAATGGAGCCACTGGCCAC 2473
Db |||||
QY 2579 CGAAGGCCAGTGGGTAGCGTGTGACGACCGCTGGGAATCTGGCTGGAGACTGGTGT 2638
Db |||||
QY 2474 AGCAGAGTTGGATGCGCGGAGGTACTCCCAATGGCAGCAGTGAACAAAGCCCTGAGA 2533
Db |||||
QY 2639 CTGACAGAGCTTGGATACAAAGTGTTCAAAGTGTGCATTAAGCGAGCTTATTTGAAA 2698
Db |||||
QY 2534 GAGGCTGGCATGAGTTGAATCTGGGTACCGGAGAGCAT 2573
Db |||||
QY 2699 AGGTACGGGTCCAATATGGCTGAATGAAGTATTTTGTTCGGGAAAGTCAATCATGA 2758
Db |||||
QY 2574 --TCAGTGATCTAACATCCAGCTTCATATAACCCAGGACAGTATATCAGAGCTTTGA 2631
Db |||||
QY 2759 AGAGTCAGATTAGACAGTGGGTGTGAGAGCTGTTCGACACAGAGATGCTGGGG 2818
Db |||||
QY 2632 ACAGGTAGTGAATGAATCTTTTCGGGATGGGTAAACTGGGCTGCGATTTGGGCTCTC 2691
Db |||||
QY 2819 TCACCTTGACCTACATAATGCATCATATTTTCAATTTTCAATTTTAACTGTTTAAAG 2878
Db |||||
QY 2692 CTCCTTTGGGGGGCAGTGTGCTGGAAAGCGTGAACAGAGATGCAAGGTATTGGTGAG 2751
Db |||||
QY 2879 TGATTTTTTCTTTGCTTCACTAAATCAGCTTAAATTAATTAATTAAGAAACTAAGAAAT 2938
Db |||||
QY 2752 TCGGATTCAGATTTGGATGGCCACCTACCTGAATGACCAAC--TAGAGCTTTGGATCCAGG 2810
Db |||||
QY 2939 TTATCCAGAAAGGAATATTTAAATATCTCGGATAAACAATATAAATAGCTTCATAT 2998
Db |||||
QY 2811 AGAACGGCGGTGGGACACTTTTGTGGATCTCTACGGGAAACAATACAGCAACCCAGAG-- 2867
Db |||||
QY 2999 TTGCTTCAATACAGAACCAATTTCACTTCTAGAGTTTAAAGTGGCTGCTGCCGAT 3058
Db |||||
QY 2868 --GCCGAAAGGCCAGAGCGTTTCAACCGCT---GGTTCCTGACGGGCACTGCTGGC 2922
Db |||||
QY 3059 TGATCCCTCAGGATATAGTATTTGCTTTGCTATAGGAGGGGAAATGATGCTTAT 3118
Db |||||
QY 2923 TGGTGTAGTTCTGCT---GGGCTCACTCTTCAGTCCGGAAGTACCAGATTTCTGAGATAT 2979
Db |||||
QY 3119 GCAATATCTTTGTAGTCTTGAACATGATGAACGATAGTTAGCAACATGCTTAAAGGA 3178
Db |||||
QY 2980 CCATCACACTGGCGCGGAAATCCGGTCGAGTCTAGAGGGCCAGATCTAGATCGATAC 3039
Db |||||
QY 3179 GAGAAAGCACCGTGCATCCGATTTGGTGAAGTGAAGTGGTGGTGCATGCTGCTTATTA 3238
Db |||||
QY 3040 GTCACATATGATCGGCTGTGGATGTGTGAGTTAGGGGTGGAAAGTCCCGAGGCTC 3099
Db |||||

QY 3239 GGAAGGCCAACAGACGGGTCTGACATGATTGGAGAAACCACTGAATTCGCAATTCACAG 3298
Db |||||
QY 3100 CCGAGCAGGCAGA----- 3112
Db |||||
QY 3299 ATATTGTATTAAAGTGTGCTAGCTCGATACAGAAACGCCATTTTGACCAATTCACACATTG 3358
Db |||||
QY 3113 -----AGTATGCAAGCATGC 3128
Db |||||
QY 3359 GTGTGCACCTCAAGCTTACGCTGCGCAAGCACTCAGGCGCAAGGGCTGTCTAAAGGA 3418
Db |||||
QY 3129 ATCTCAATTAGTCAGCAACACAGGTGTGMAAGTCCCGAGGCTCCCGACGACGAGNAGTA 3188
Db |||||
QY 3419 AGCGGAACACGTAGAAAGCCAGTCCGAGAAACCGTGTGTGACCCCGGATGAATGTACGT 3478
Db |||||
QY 3189 TGCAAGC-----ATGCATCTCAATT 3209
Db |||||
QY 3479 ACTGGGCTATCTGCAAGGGAACCAAGCGCAAGAGAAAGACAGGTAGCTTTCAGTG 3538
Db |||||
QY 3210 AGTCAGCAACCATAGTCCGCCCCCTAACTCCGCCATCCCGCCCCCTAACTCCGCCAGTT 3269
Db |||||
QY 3539 GGCTTACATGGCATAGCTAGACTGGCGGTGTTTATGACAGCAAGCAACCGGAATTCG 3598
Db |||||
QY 3270 CCGCCCATTTCCGCCCATGGCTGACTAATTTTTTTTATTTATGACAGAGCCGA----- 3324
Db |||||
QY 3599 CAGCTGGGGCGCTCTGTGTAAAGTTGGGAAGCCCTGCAAG--TAAACTGGATGGCTTT 3656
Db |||||
QY 3325 -GGCGGCTCGGCTCTGAGCTATTCAGAAAGTAGTAGAGGCTTTTTTGGAGGCTTAG 3383
Db |||||
QY 3657 CTTGCCGCCAAGGATCTGATGGCGCAGGGATCAAGATCTGATCAAGAGACAGATGAGG 3716
Db |||||
QY 3384 GCTTTTGCAAAAGCTTGGGCTGCAGGTCGAGGCGGATCTGATCAAGAGACAGATGAGG 3443
Db |||||
QY 3717 ATCGTTTCGATGATTGAACAAAGATGATTCAGCGCAGGTTCTCGGCCGCTTGGGTGGA 3776
Db |||||
QY 3444 ATCGTTTCGATGATTGAACAAAGATGATTCAGCGCAGGTTCTCGGCCGCTTGGGTGGA 3503
Db |||||
QY 3777 GAGCTATTGGCTATGACTGGGCAACAGACAAATCGGCTGCTCTGATGCCGCCGTGTT 3836
Db |||||
QY 3504 GAGGCTATTGGCTATGACTGGGCAACAGACAAATCGGCTGCTCTGATGCCCGGTGTT 3563
Db |||||
QY 3837 CCGGCTCTGACGCGAGGGGCGCGGTTCTTTTGTCAAGACCACTGTTCGGTGCCT 3896
Db |||||
QY 3564 CCGGCTCTGACGCGAGGGGCGCGGTTCTTTTGTCAAGACCACTGTTCGGTGCCT 3623
Db |||||
QY 3897 GAATGAATGACGACGAGCGCGCTATCGTGTGCTGCGCACGACGCGGCTTCTCTTG 3956
Db |||||
QY 3624 GAATGAATGACGACGAGCGCGCTATCGTGTGCTGCGCACGACGCGGCTTCTCTTG 3683
Db |||||
QY 3957 CGCAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGACTGGCTGCTATTGGGCGAAGT 4016
Db |||||
QY 3684 CGCAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGACTGGCTGCTATTGGGCGAAGT 3743
Db |||||
QY 4017 GCGGGGCGAGATCTCTGTCTATCTACCTGTCTCTGCGCGAGAAAGTATCCATCATGGC 4076
Db |||||
QY 3744 GCGGGGCGAGATCTCTGTCTATCTACCTGTCTCTGCGCGAGAAAGTATCCATCATGGC 3803
Db |||||
QY 4077 TGATGCAATCGGCGGCTGCATAGCTTGTGATCGGCTACCTGCGCATTTCAACCAACCAAGC 4136
Db |||||
QY 3804 TGATGCAATCGGCGGCTGCATAGCTTGTGATCGGCTACCTGCGCATTCGACCAACCAAGC 3863
Db |||||
QY 4137 GAAACATCGCATCGAGCGACGATCTCGGATGGAAGCGGCTTGTTCGATCAGGATGA 4196
Db |||||
QY 3864 GAAACATCGCATCGAGCGACGATCTCGGATGGAAGCGGCTTGTTCGATCAGGATGA 3923
Db |||||
QY 4197 TCTGACGAAGAGATCAGGGGCTCGCGCAGCGGAATCTGTTGCGCAGGCTCAAGGCGG 4256
Db |||||
QY 3924 TCTGACGAAGAGATCAGGGGCTCGCGCAGCGGAATCTGTTGCGCAGGCTCAAGGCGG 3983
Db |||||
QY 4257 CATGCCGAGCGGAGATCTCGTGTGACCCATGGCGATGCCCTGCTTCCGATATCAT 4316
Db |||||
QY 3984 CATGCCGAGCGGAGATCTCGTGTGACCCATGGCGATGCCCTGCTTCCGATATCAT 4043
Db |||||
QY 4317 GGTGGAATAATGGCGCTTTTCTGATTCATCGACTGTGGCGGCTGGGTGTGGCGGACCG 4376
Db |||||

Db 4044 GGTGAAATGGCGCTTTCTGATTCATCGACTGTGCGCGCTGGGTGCGGACG 4103
Qy 4377 CTATCAGACATAGCGTGGCTACCCGATGATTTGCTGAAGAGCTTGGCGCGATGGGC 4436
Db 4104 CTATCAGACATAGCGTGGCTACCCGATGATTTGCTGAAGAGCTTGGCGCGATGGGC 4163
Qy 4437 TGACCGCTTCTCGTGTCTTTACGGTATCGCGCTCCCGATTCGAGCGCATCGCCTCTTA 4496
Db 4164 TGACCGCTTCTCGTGTCTTTACGGTATCGCGCTCCCGATTCGAGCGCATCGCCTCTTA 4223
Qy 4497 TCGCCTTCTGACGAGTTTCTTCTGAGCGGGACTCTGGGTTTCGATAAATAAAGATTTT 4556
Db 4224 TCGCCTTCTGACGAGTTTCTTCTGAGCGGGACTCTGGGTTTCGATAAATAAAGATTTT 4283
Qy 4557 ATTTAGTCTCCAGAAAAGGGGGATCAAGAGACCCACCTGTAGTTTGCAGAGCTAGC 4616
Db 4284 ATTTAGTCTCCAGAAAAGGGGGATCAAGAGACCCACCTGTAGTTTGCAGAGCTAGC 4343
Qy 4617 TTAAGTAACGCCATTTTCAAGGCAATGAAAATACATAACTGAGATAGAGAAGTTTCAG 4676
Db 4344 TTAAGTAACGCCATTTTCAAGGCAATGAAAATACATAACTGAGATAGAGAAGTTTCAG 4403
Qy 4677 ATCAAGGTCAGGAACAGATGGAACAGCTGAATATGGGCCAAACAGGATATCTGTGTAAG 4736
Db 4404 ATCAAGGTCAGGAACAGATGGAACAGCTGAATATGGGCCAAACAGGATATCTGTGTAAG 4463
Qy 4737 CAGTTCTTCCCGCGCTCAGGCGCCAGAACAGATGGAACAGCTGAATATGGGCCAAACAG 4796
Db 4464 CAGTTCTTCCCGCGCTCAGGCGCCAGAACAGATGGAACAGCTGAATATGGGCCAAACAG 4523
Qy 4797 GATATCTGTGTAGCAGTTCTCTCGCGGCTCAGGCGCCAGAACAGATGTTCCCAAGGA 4856
Db 4524 GATATCTGTGTAGCAGTTCTCTCGCGGCTCAGGCGCCAGAACAGATGTTCCCAAGGA 4583
Qy 4857 CCGGTCCAGCCCTCAGCAGTTTCTAGAGAACCATCAGATGTTTCCAGGGTCCCAAGGA 4916
Db 4584 CCGGTCCAGCCCTCAGCAGTTTCTAGAGAACCATCAGATGTTTCCAGGGTCCCAAGGA 4643
Qy 4917 CCTGAATGACCTGTGCTTATTTGAACCAATCAGTTCCGTTCTCGTTCTGTTTC 4976
Db 4644 CCTGAATGACCTGTGCTTATTTGAACCAATCAGTTCCGTTCTCGTTCTGTTTC 4703
Qy 4977 GCGGCTTCTGCTCCCGAGCTCAATAAAGAGCCCAACCCCTCACTCGGGGCGCCAG 5036
Db 4704 GCGGCTTCTGCTCCCGAGCTCAATAAAGAGCCCAACCCCTCACTCGGGGCGCCAG 4763
Qy 5037 TCCTCCGATTGACTGAGTGCCTCGCGGTACCGGTATCCCAATAAACCCCTTTGCAAGTTGC 5096
Db 4764 TCCTCCGATTGACTGAGTGCCTCGCGGTACCGGTATCCCAATAAACCCCTTTGCAAGTTGC 4823
Qy 5097 ATCCGACTTGTGTTCTGCTGTTCTGGGAGGCTCTCTCTGAGTGAATGACTACCGGT 5156
Db 4824 ATCCGACTTGTGTTCTGCTGTTCTGGGAGGCTCTCTCTGAGTGAATGACTACCGGT 4883
Qy 5157 CAGCGGGGCTTTCATTTGG 5177
Db 4884 CAGCGGGGCTTTCATTTGG 4904

RESULT 4
ID AAX90485
XX AAX90485 standard; DNA; 7546 BP.
AC AAX90485;
XX
XX
DT 29-SEP-1999 (first entry)
XX
DE Plasmid retroviral vector pLXSU-IR5-N nucleotide sequence.
KW Plasmid retroviral vector; expression system; immunogenic; gene therapy;
KW immune response; immunosuppression; gene delivery; therapeutic;
KW MHC-I autoimmune disease; tumour; 88.

XX Synthetic.
OS
XX WO9936562-A1.
PN
XX 22-JUL-1999.
PD
XX 13-JAN-1999; 99WO-US000733.
PF
XX 14-JAN-1998; 98US-0071409P.
PR
XX (HUMA-) HUMAN GENE THERAPY RES INST.
PA
XX Radosevich TJ, Link CJ;
XX WPI; 1999-468988/39.
XX
XX Expression system containing therapeutic gene and an immunosuppressor
XX gene useful for treating an MHC-I autoimmune disease or killing tumor
XX cells.
XX
XX Disclosure; Page 147-151; 154pp; English.
XX
XX The present invention describes a nucleotide expression system for the
XX introduction of a therapeutic gene comprising: (i) a nucleotide sequence
XX encoding an immune suppression gene; (ii) a promoter; and (iii) a
XX transcription termination signal, where the system is able to inhibit,
XX evade or eliminate a recipient cell immune response to the therapeutic
XX gene when the gene is transformed into a recipient cell. AAX90481 to
XX AAX90484 represent specifically claimed plasmid retroviral vector
XX nucleotide sequences from the present invention. The expression system
XX and vectors containing it can be used for gene therapy, for treating an
XX MHC-I autoimmune disease or for killing tumor cells. The expression
XX system contains an immunosuppressive gene which prevents host rejection
XX of the vector. The present sequence represents a plasmid retroviral
XX vector from the present invention, designated pLXSU-IR5-N
XX
SQ Sequence 7546 BP; 1695 A; 2066 C; 1981 G; 1804 T; 0 U; 0 Other;
Query Match 37.08; Score 1916.6; DB 2; Length 7546;
Best Local Similarity 64.24; Pred. No. 0;
Matches 3348; Conservative 0; Mismatches 1624; Indels 247; Gaps 21;
Qy 1 TTTGAAAGACCCACCGTAGGTGGCAAGCTAGCTTAAGTAACGCCACTTTGCAAGGCAT 60
Db 175 TTTGAAAGACCCACCGTAGGTGGCAAGCTAGCTTAAGTAACGCCACTTTGCAAGGCAT 234
Qy 61 GGAAAAATACATAAATCTGAGAAATAGAAAAGTTTCAGATCAAGGTCAAGAAACAAAGAACAGC 120
Db 235 GGAAAAATACATAAATCTGAGAAATAGAAAAGTTTCAGATCAAGGTCAAGAAACAAAGAACAGC 294
Qy 121 TGAATACCAACAGGATATCTGTGGTAAGCGGTTCTGCCCCGGCTCAGGGCCAGAAACA 180
Db 295 TGAATACCAACAGGATATCTGTGGTAAGCGGTTCTGCCCCGGCTCAGGGCCAGAAACA 354
Qy 181 GATGAGACAGCTGAGTATGCGGCCAAGACAGATATCTGTGTAAGAGCTTCTCCCGCGG 240
Db 355 GATGAGACAGCTGAGTATGCGGCCAAGACAGATATCTGTGTAAGAGCTTCTCCCGCGG 414
Qy 241 CTCGGGGCCCAAGAACAGATGTTCCCGAGATGCGGTCCAGGCCCTCAGCAGTTTCTAGTGAA 300
Db 415 CTCGGGGCCCAAGAACAGATGTTCCCGAGATGCGGTCCAGGCCCTCAGCAGTTTCTAGTGAA 474
Qy 301 TCATCAGATGTTTCAGGGTGGCCCAAGGACCTGAAATGACCCCTGTACCTTATTTGAAAC 360
Db 475 TCATCAGATGTTTCAGGGTGGCCCAAGGACCTGAAATGACCCCTGTACCTTATTTGAAAC 534
Qy 361 TAACCAATCAGTTCCGCTTCTCGCTTCTGCGGCGCTTCCGCTCTCCGAGCTCAATAAA 420
Db 535 TAACCAATCAGTTCCGCTTCTCGCTTCTGCGGCGCTTCCGCTCTCCGAGCTCAATAAA 594
Qy 421 AGAGCCCAACACCCCTCACTCGCGCGCCCAAGTCTTCCGATAGACTGCGTCCCGGGGTAC 480

| | | |
|------|--|------|
| 595 | AGAGCCCAACCCCTCACTCGGGCGCGCAGCTCTCTCCGATAGACTCGGTGCGCCCGGGTAC | 654 |
| 481 | CCGTATTCCCAATAAAGACCTCTTGCTGTTTGCATCCGAATCGTGGTCTCGCTGTCTCCTTG | 540 |
| 655 | CCGTATTCCCAATAAAGACCTCTTGCTGTTTGCATCCGAATCGTGGTCTCGCTGTCTCCTTG | 714 |
| 541 | GGAGGGTCTCTCTGAGTGATGACTACCAAGACGGGGGTCTTTCAATTTGGGGGGCTCGT | 600 |
| 715 | GGAGGGTCTCTCTGAGTGATGACTACCAAGACGGGGGTCTTTCAATTTGGGGGGCTCGT | 774 |
| 601 | CCGGGATTTTGAGACCCCTCGCCAGGGACCAACCCACCCAGCCGGGAGGTAAAGCTGCGCC | 660 |
| 775 | CCGGGATTTTGAGACCCCTCGCCAGGGACCAACCCACCCAGCCGGGAGGTAAAGCTGCGCC | 834 |
| 661 | AGCAACTTATCTGTCTGTCTGCCGATTGTCTAGTGTCTATGTTTGATGTATGCGCCTCGC | 720 |
| 835 | AGCAACTTATCTGTCTGTCTGCCGATTGTCTAGTGTCTATGTTTGATGTATGCGCCTCGC | 894 |
| 721 | TCTGTACTAGTTAGTTACTAGTCTCTGTATCTGGGGACCCGTCGTGTGAACTGACGAGTT | 780 |
| 895 | TCTGTACTAGTTAGTTACTAGTCTCTGTATCTGGCGGACCCGTCGTGTGAACTGACGAGTT | 954 |
| 781 | CTGAACACCCGGCGCACCTCTGGGAGACGTCGCCAGGACCTTTGGGGGGCGTTTTTGTGG | 840 |
| 955 | CTGAACACCCGGCGCACCTCTGGGAGACGTCGCCAGGACCTTTGGGGGGCGTTTTTGTGG | 1014 |
| 841 | CCCGACCTGAGAAAGGAGTCCGATGTGGAATCCGACCCCGTCAGGATATGTGGTCTTGGT | 900 |
| 1015 | CCCGACCTGAGAAAGGAGTCCGATGTGGAATCCGACCCCGTCAGGATATGTGGTCTTGGT | 1074 |
| 901 | AGGAGACGAGAACCTAAACAGTTCCCGCCTCCGTCCTGAAATTTTGTCTTGGTTTGGAA | 960 |
| 1075 | AGGAGACGAGAACCTAAACAGTTCCCGCCTCCGTCCTGAAATTTTGTCTTGGTTTGGAA | 1134 |
| 961 | CCGAGCGCGCGCTTGTCTGCTGAGCCAGCAAGCTTGGGCTGCAGGTCGACTAGAGGA | 1020 |
| 1135 | CCGAGCGCGCGCTTGTCTGCTGAG----- | 1162 |
| 1021 | TCAAATCGGCACGAGTAAATCGGTGCTGCGCTCTTTAGGACATATGAAGTATGGCACAGT | 1080 |
| 1163 | -----CGCTGCAGCATCGTCTCTGTGTGTCTGTCTGCTGACTGTGT | 1202 |
| 1081 | GGGATGACTTTCGTGATCAGCAAGAGACACTGACAGCTGTGACAGAGTCTGTGAAGTTCG | 1140 |
| 1203 | TTCTGTATTGTCTGAAAAATTAGGGCCAGACTGTTACC-----ACTCCCTAAAGTTTG | 1255 |
| 1141 | ATGCTCGCTCAGTGACAGCTTTGTTCTCTCCCACTCTAAAAATGGCGCCAACTCTCAAG | 1200 |
| 1256 | ACCTTAGGTCACTGGAAAGATGTGCGAGCGGATCGCTCACACCAAGTCGGTAGATGTCAG | 1315 |
| 1201 | AGAGGATGAAGTCTTTATAAACTGCACCTGATCACCTTTATCTCATTTGTGTTCTAGTTC | 1260 |
| 1316 | AAGAGAGTTGGGTTA-CCTTCTGCTCTGCGAATGGCCAACTTTTAACGTCGGATGGCC | 1374 |
| 1261 | TCGTGCCCATCANTTGGCATAGTGGCAGTCAAGTCTCTGAAATGGGAAACGAAAGATTGCA | 1320 |
| 1375 | GCGAGACGGCACCTTTAAACCGAGACCTCATCACCCAGGTTTAAGATCAAGGTCTTTTCA | 1434 |
| 1321 | CGGTTGGTCTAGTTAATGAGATATATCTCCAAGTCCGGAAGGCAAGGAAATGGCAGTG | 1380 |
| 1435 | TGGCCCGCATGGACA-----CCGAGACCAAGTCCCTCATCATCGTGACCTCG | 1480 |
| 1381 | AAGATGAATGAGATTTGAGAACTGTGATGGAAACCATGAGCAACATGGAAGACGAA | 1440 |
| 1481 | GAAGCCTTGGCTTTTGACCCCGCTCTCTCGGTGCAAGCCCTTTGACACCTTAAGCCTCG | 1540 |
| 1441 | TCCAGTATCTTTACAGATAATGAAGCCAACTCTCTAGATGCTGAAGATTTTCCAAATTTCA | 1500 |
| 1541 | CCTCCTCTTCCATCCGCGCCGCTCTCTCCCGCTTGAACTCTCTCGTTCGACCCCGCCT | 1600 |
| 1501 | GCATAACAACCTGATCAAGAATTTAATGATGTCTTTTCCAGTAAATTCCTACTTTCTCT | 1560 |
| 1601 | CGATCTCTCTTTATCCAGCCCTCACTCTCTCTAGGCGCGGAAATTCGTTACTCGAG | 1660 |

| | | | |
|----|------|--|------|
| Qy | 1561 | CCATCCAGGAACATGAGAAATATCATATAGGGGATATCTCCAAGTCAATTAGTAGGTCTGAACA | 1620 |
| Db | 1561 | GATCCGGCTGTGGAATGTGTGTAGTTAGGGTGTGAAAGTCCCCAGGCTCCCAAGCAGG | 1720 |
| Qy | 1621 | CCACAGTACTTGTATTTGTCAGTTCAGTATTGAAACACTGAATGCGAGAGTCCAAGAGAAATG | 1680 |
| Db | 1721 | CAGNAGTA-----TGCAAAGCATGCAATCTCAATTAGTACAGCAACAGGTGTGGAAGT | 1773 |
| Qy | 1681 | CATTATAAACAAAGAGGAGATCGTAAATTTAGAGGAGCGTATATACAATGCATCAGCAG | 1740 |
| Db | 1774 | CCCCAGGCTCCCCAGCAGGAGGAAGTATGCAAAAGCATGCATCTCAATTAGTCAGCAACCA | 1833 |
| Qy | 1741 | AAATTAAAGTCTCTAGATGAAAAACAAGTATATTTTGGAAACAGGAATAATAAAGGGGAATGA | 1800 |
| Db | 1834 | TAGTCCCGCCCTTAACTTCC--GCCCATCCCGCCCTTAACTTCCGCCCAAGTTCCGCCCAATC | 1891 |
| Qy | 1801 | AACGTGTTGAATATATCACTAATGATCTGAGGCTGAAGGATTTGGGAACATTTCTCAGACAT | 1860 |
| Db | 1892 | TCGCCCCCATGGCTGACTAAATTTTTTTTATTATTAAGCAGAGCGCAGCGCCCTCGGCCCT | 1951 |
| Qy | 1861 | TGAAAAATATCACTTTACTCCAAGGTGCCAAGAAATGCTCGCTGACTGGGAAATGACCA | 1920 |
| Db | 1952 | TGAGCTA-----TTCCAGAGTAGTAGGAGGCTTTTTCGAGGCGCT | 1993 |
| Qy | 1921 | ACGATCTGGGCTCCAAATGACCATCGGGCTGTGAACAGCAGAGAGTGAAATTCACAGGCA | 1980 |
| Db | 1994 | AGCGTTTTCGAAAAAGCTTGGGCTGCAAGCTTGGTACCGAGCTCCGATCGATATCTGCGG | 2053 |
| Qy | 1981 | CCTACATCAGCGGTAAACAGCCACATCAAAATGAGATCAAAAGATCACCACCTGCATGGGA | 2040 |
| Db | 2054 | CCGCGTCGACGGATGAACCTTGTAAATGCTTATTTCTAGCCCTCTGGGCCCGCGTCCGGGT | 2113 |
| Qy | 2041 | CACAAAACACATCAACACAGAGAGCACCAGCCCACTTTGGCTTTCACCGTCAATTGGAAGT | 2100 |
| Db | 2114 | AGTATGCCTGAAT-----TATCTGTGACTCTTTTCGATGAACCTCGGCCCTTGGTG | 2164 |
| Qy | 2101 | TTTTCAGAGTCCACCACTGTCTTCACGGGCCGAGTGTTCATAGACAGGAATGGGAAGGAG | 2160 |
| Db | 2165 | GAGACGGAGCGGTTACCGCCTCTGTCGGATGTTTCGGAGTACCGAGTAGAGTATTCGAG | 2224 |
| Qy | 2161 | TCCTGAAGACCATGTGGCTGCTCGGTCAAGTGTTAATGCACATTTGTGTGATGACTGGAAG | 2220 |
| Db | 2225 | GCGCGCTGCGTGTCTCGATCGGCGGTGCACTGGAGGCTCTGTGAGCCCTCGCGCGGAAC | 2284 |
| Qy | 2221 | CTACCGGCTCGGCATCAACATCTTCACTCGCTCGCACACAGAAAGGAGTGCAGTGC | 2280 |
| Db | 2285 | CTGTCGTGCCACGCCGACA-----CCCCGGGTGTACTACAGACGCTGGAGGGCT | 2336 |
| Qy | 2281 | ACCAAGGTCTCTCTGGAATCCAGGTGAAAAAGAGATAGAGGCCCTCTCTGGACAAATGG | 2340 |
| Db | 2337 | ACGGGATCGAGTGCAGCGCGGTGGAGGAGCTCTCCGAAAGCCTCGTCGCAAA----- | 2391 |
| Qy | 2341 | TATACAGCGTTTTCAGGTCTAATAGGTACTCCAGTCTTTAAAGTGTGATCGGGGGATCT | 2400 |
| Db | 2392 | -AGCTACTGGCTCCGGGACTATCGTGTTCCTCCCAACGCCACAAAATCTCGTGTGTTCTACT | 2450 |
| Qy | 2401 | CT--GGTTTACCTGGAGTTCGAGGATTCACAGAACCAATGGGGAGACCCGGNAGCCAGG | 2458 |
| Db | 2451 | TTTTCCCGCTGCCACCAATGCCAACTTATATGTAGATGCGGAACCCCGGTGCTCGTGC | 2510 |
| Qy | 2459 | ACTTAATGGACAAAAAGGCCAGAAAGGAGTGGGAAGCATGCAAAAGACAAATC | 2518 |
| Db | 2511 | CTTGGGTTCCCTGTGGAGTCTGTTAGAGACATCGAACGACTATTTGTTTGAAGATCGCC | 2570 |
| Qy | 2519 | TAATACAGTCCGACTGGTGGGTGGCAGCGGCCCTCACGAAGGCAGAGTGAGATTTTTCA | 2578 |
| Db | 2571 | GTCTAATGGCGTACT-----ACGCGCTCACGATTAAAGTCGCGCGCATATACGCTGATGA | 2624 |
| Qy | 2579 | CGAAGGCCAGTGGGTACCGTGTGTGACGACCGCTGGGAACTCGCTGTGGAGGACTGGTCTG | 2638 |
| Db | 2625 | TGCTGGCAGTGAATTCAGTGTGTTTTGGGGGCTGTATGTGAAAGGTTGGCTGTCACCCGACAT | 2684 |


```

Db 4770 CCAAGACAGATGGAACAGCTGAATATGGGCCAAACAGAGATATCTGTGTAAGCAGTTCC 4829
QY 4819 TCCTCCCGGCTCAGGCGCAAGAACAGATGTCCTCCAGATGCGGTCCAGCCCTCAGCAGTTT 4878
Db 4830 TCCTCCCGGCTCAGGCGCAAGAACAGATGTCCTCCAGATGCGGTCCAGCCCTCAGCAGTTT 4889
QY 4879 CTAGAGAACCATCAGATGTTTCCAGGTTGCCAGGTCGCCACAGGACCTGAATGACCTGTGCTTA 4938
Db 4890 CTAGAGAACCATCAGATGTTTCCAGGTTGCCAGGTCGCCACAGGACCTGAATGACCTGTGCTTA 4949
QY 4939 TTTGAACCTAACCAATCAGTTGCTTCTCGTCTCTGCTTCTGCTGCGGCTTCTGCTCCCGAGCT 4998
Db 4950 TTTGAACCTAACCAATCAGTTGCTTCTGCTTCTGCTTCTGCTGCGGCTTCTGCTCCCGAGCT 5009
QY 4999 CAATAAAGAGGCCCAACCCCTCACTCGGGCGCCAGTCTCTCGATTGACTGAGTCGCC 5058
Db 5010 CAATAAAGAGGCCCAACCCCTCACTCGGGCGCCAGTCTCTCGATTGACTGAGTCGCC 5069
QY 5059 CGGGTACCGGTATCCCAATAAACCCCTTTCGAGTTGATCCGACTTGTGCTGCTGT 5118
Db 5070 CGGGTACCGGTATCCCAATAAACCCCTTTCGAGTTGATCCGACTTGTGCTGCTGT 5129
QY 5119 TCCTTGGAGGCTCTCTCTGATGATGACTACCCGTCAGCGGGGCTTTCATTGG 5177
Db 5130 TCCTTGGAGGCTCTCTCTGATGATGACTACCCGTCAGCGGGGCTTTCATTGG 5188

RESULT 5
AAQ74450
ID AAQ74450 standard; DNA; 8316 BP.
XX
AC AAQ74450;
XX
DT 22-MAY-1995 (first entry)
XX
DE myoD retroviral expression vector pLHMDN-NSA.
XX
KW MyoD; muscle; gene expression; myoblast; muscle differentiation;
KW rhabdomyosarcoma cell; nuclear regulatory factor; chemotherapeutic agent;
KW identification; detection; skeletal muscle; ss.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT LTR 1..1159
FT FT /*tag= a
FT misc_feature 1160..1640
FT FT /*tag= b
FT FT /label= Psi region.
FT CDS 1641..2928
FT FT /*tag= c
FT FT /product= "Histidinol dehydrogenase."
FT misc_feature 2929..4389
FT FT /*tag= d
FT FT /label= MyoD NSA ApaI fragment
FT FT /note= "Comprises the myoD proximal and distal regulatory
FT FT regions."
FT CDS 4390..5259
FT FT /*tag= e
FT FT /product= "Neomycin phosphotransferase."
FT LTR 5260..5964
FT FT /*tag= f
XX
US5352595-A.
XX
PD 04-OCT-1994.
XX
PF 03-SEP-1991; 91US-00753520.
XX
PR 03-SEP-1991; 91US-00753520.
XX
PA (HUTC-) HUTCHINSON CANCER RES CENT FRED.
XX

```

```

PI Palmer TD, Tapscott SJ, Weintraub HM;
XX WPI; 1994-316186/39.
XX
PT Isolated proximal and distal myoD regulatory regions - and vectors contg.
PT them, for tissue selective gene expression in muscle cells and for
XX positive or negative muscle cell selection.
XX Example 8; Fig 8; 48pp; English.
XX
CC The myoD proximal and distal regulatory regions (See AA074447, AA074448)
CC may be inserted into vectors such as this and used (1) to induce a muscle
CC phenotype in a non-muscle cell, (2) for positive or negative selection of
CC muscle cells (3) for targeted gene expression, specifically in skeletal
CC muscle, (4) for the expression of genes controlling growth of myoblasts
CC or mutant forms of myoD that modulate muscle differentiation, (5) for
CC engineering rhabdomyosarcoma cells to increase sensitivity to
CC chemotherapeutic agents, (6) for the expression of growth factors to
CC stimulate skeletal muscle growth in animals, and (7) to identify nuclear
CC regulatory factors that bind to the regulatory region
XX
SQ Sequence 8316 BP; 1922 A; 2246 C; 2255 G; 1880 T; 0 U; 13 Other;
Query Match 32.1%; Score 1661.6; DB 2; Length 8316;
Best Local Similarity 61.1%; Pred. No. 0;
Matches 3536; Conservative 0; Mismatches 1641; Indels 607; Gaps 29;
QY 1 TTTGAAAGACCCCAACCCGCTAGGTGGCAAGCTAGCTTAACTAAGTAAACGCCACTTTGCAAGGCAT 60
Db 175 TTTGAAAGACCCCAACCCGCTAGGTGGCAAGCTAGCTTAACTAAGTAAACGCCACTTTGCAAGGCAT 234
QY 61 GGAATAATACATACTCAGTAATAGCAAAAGTTTCAGATCAAGCTCAGGAACAAGAAACAGC 120
Db 235 GGAATAATACATACTCAGTAATAGCAAAAGTTTCAGATCAAGCTCAGGAACAAGAAACAGC 294
QY 121 TGAATACCAACAGCATATCTGTGTAAGCGGTTCTGCCCCGGCTCAGGCCCAAGAAC 180
Db 295 TGAATACCAACAGCATATCTGTGTAAGCGGTTCTGCCCCGGCTCAGGCCCAAGAAC 354
QY 181 GATCAGACAGCTGATGATGGCCCAACAGCATATCTGTGTAAGCGGTTCTGCCCCGG 240
Db 355 GATCAGACAGCTGATGATGGCCCAACAGCATATCTGTGTAAGCGGTTCTGCCCCGG 414
QY 241 CTCGGGGCCCAAGAACAGATGTCCTCCAGATGCGGTCCAGCCCTCAGCAGTTTCTAGTGA 300
Db 415 CTCGGGGCCCAAGAACAGATGTCCTCCAGATGCGGTCCAGCCCTCAGCAGTTTCTAGTGA 474
QY 301 TCATCAGATGTTTCCAGGGTGGCCCAAGGACCTGAAATGACCCCTGACCTTATTTGAAC 360
Db 475 TCATCAGATGTTTCCAGGGTGGCCCAAGGACCTGAAATGACCCCTGACCTTATTTGAAC 534
QY 361 TAACCAATCAGTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCT 420
Db 535 TAACCAATCAGTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCT 594
QY 421 AGAGCCCAACACCCCTCAGTCCGGCGCCAGTCTTCCGATAGACTGCGTCCCGGGGTAC 480
Db 595 AGAGCCCAACACCCCTCAGTCCGGCGCCAGTCTTCCGATAGACTGCGTCCCGGGGTAC 654
QY 481 CCGTATTTCCCAATAAAGCCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCT 540
Db 655 CCGTATTTCCCAATAAAGCCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCT 714
QY 541 GGAGGGTCTCTCTGAGTGATGACTACCAACGAGGGGGTCTTTCATTTGGGGGCTCGT 600
Db 715 GGAGGGTCTCTCTGAGTGATGACTACCAACGAGGGGGTCTTTCATTTGGGGGCTCGT 774
QY 601 CCGGGAATTTGGAGACCCCTGCGCAGGAGCACCCAGCCACACCCAGCCAGGAGGTAAAGTGGCC 660
Db 775 CCGGGAATTTGGAGACCCCTGCGCAGGAGCACCCAGCCACACCCAGCCAGGAGGTAAAGTGGCC 834
QY 661 AGCAACTTATCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720

```

Db 835 AGCAACTTATCTGTCTGTCCGATTGTCTAGTGTCTATGTTTGTATGCTTATGCGCTGCG 894
Qy 721 TCTGTACTAGTAACTAGTCTGTATCTGGCGGACCCGTGTGTGAACTGACGAGTT 780
Db 895 TCTGTACTAGTAACTAGTCTGTATCTGGCGGACCCGTGTGTGAACTGACGAGTT 954
Qy 781 CTGAACACCCGCGCCCAACCTCTGGGAGACGTCCACAGGACCTTTGGGGCCGCTTTTGTGG 840
Db 955 CTGAACACCCGCGCCCAACCTCTGGGAGACGTCCACAGGACCTTTGGGGCCGCTTTTGTGG 1014
Qy 841 CCGGACCTGAGGAAGGGAGTGCATGTGGAATCCGACCCCGCTCAGGATATGGTTCTGGT 900
Db 1015 CCGGACCTGAGGAAGGGAGTGCATGTGGAATCCGACCCCGCTCAGGATATGGTTCTGGT 1074
Qy 901 AGGAGACGAGAACCTTAAACAGTTCCCGCTCCGCTCTGAAATTTTGTCTTTCGGTTGGAA 960
Db 1075 AGGAGACGAGAACCTTAAACAGTTCCCGCTCCGCTCTGAAATTTTGTCTTTCGGTTGGAA 1134
Qy 961 CCGAAGCCGCGCTCTGTCTGTCTGCTGACGCAAGCTTTGGGCTGCGAGTCTCTAGAGGA 1020
Db 1135 CCGAAGCCGCGCTCTGTCTGTCTGCTGACGATCGTTCTGTGTGTCTCTGTCTGACTGTGT 1194
Qy 1021 TCAATTCGGCAGCAGTAATCGGTCTGCGCTCTTTTAGGACATATGAAGTATGGC----- 1075
Db 1195 TTCTGTATTTGTCTGAAATATTAGGSCCAGACTGTTTACCACCTCCCTTAAGTTTGAACCTTAG 1254
Qy 1076 -----ACAGTGGGATGACTTTCTGTATCAGCAAGAGACACTGACAGCTGTACAG 1125
Db 1255 GTCACTGGAAGATGTCGAGGGGATCGCTCAACACAGTCGCTAGATGTCGAAGAGAGAC 1314
Qy 1126 AGTCTGCAAGTTCGATCTCGCTCAGTGACAGCTTTGCTTCCCTCCCATCTTAAATG 1185
Db 1315 GTTGGTTTACCTTTCTGCTCTCGAGATGCGCAACCTTTAAGCTCGGATGCGCGGAGACG 1374
Qy 1186 GCCCAACTCTTTCAAGAGAGGATGAAGTCTTTATAAACTGCACTGATCACCCTTTATCTCA 1245
Db 1375 GCACCTTTAAACCGAGACCTCATCACCAGGTTAAGATCAAGGTCTTTTCACTTGGCCGCG 1434
Qy 1246 TTGTGTTTGTATGTTCTCTGCGCCATCATTTGGCATAGTGGCAGCTCAGCTCTGAAATGGG 1305
Db 1435 ATGGACACCCAGACCCAGGTCCCTACATCGTGACCTGGGAAGCCTTGGCTTTTGAACCCCG 1494
Qy 1306 AAACGAAGAAATTGCACCGTTGCTCAGTTAATGAGATATATCTCCAAGTCGGGAAGCA 1365
Db 1495 CTCCTCGGTCAAGCCCTTTGTACACCTAAGCCTTCGCGCTCTCTCTCCATCCGCGCC 1554
Qy 1366 AAGGAATGGCAGTGA-----AGATG 1386
Db 1555 CGTCTCTCCCTTTGAACCTCTCTGTTCCGACCCCGCTCGATCTCCCTTTATCCAGGCC 1614
Qy 1387 AAATGAGATTTGAGAACTGTATGGAAAC-----GCATGAGCAACATGGA 1432
Db 1615 TCATCTCTTCTTAGCGCCCGGAATTGATCCGGAACCATGAGCTTCAATACCTTGATTGA 1674
Qy 1433 AAGCAGAAATCAGATATTTTCAGATAATGAAGCAACTCTCTAGATGCTAAGAAATTTCCA 1492
Db 1675 CTGGACAGCTGTAGCCCTGAAACAGCAGCGGTGCGTCTGAGCGCTCCGGGATTTCCGC 1734
Qy 1493 AAATTTTCAGCATAAC-----AACTGATCAAGATTTAATGATGTTCTTTTCCAGCT 1543
Db 1735 CTCTGACAGTATTAACCGGACGCTCAGCGATATCTGGATATGTGAAAAACGCGCGGTGA 1794
Qy 1544 AAATTCCTTACTTTCTCTCCATCCAGGAACATGAGAAATCATAGGG----- 1589
Db 1795 CGATGCCCTTGGTGAATACAGCGCTTAAATTTGATAAAAAACAGAGTGACAGCGCTACGCGT 1854
Qy 1590 -----GATATCTCAAGTCAATAGTAGGTCTGAACCAACAGTACTTTGATTTGCA 1639
Db 1855 CACCCCTGAAGAGATTCGCGCGCGCGCTGAGCGAGCAATTTAAACAGGCGAT 1914
Qy 1640 GTTCAGTATTTGAAACACTGAATGGCAGAGTCCA---AGAGAATGCATTTAAACCAACAGA 1696
Db 1915 GACCGCTGCCCTCAAAAATATTGAACCGTTTCAATTCGCGCAGACGCTTACCGCTGTAGA 1974

Qy 1697 GGAGATCGTAAATTTAGAGGAGCGTATATACAAATGCATCAGCAGAAATTAAGTCTCTAGA 1756
Db 1975 TGTGAAACCCAGCCAGCGCGTGTCCAGCAGGTACGCGTCCCGTCTCGTCTGTGCG 2034
Qy 1757 T---GAAAAACAAGTATATTTGGAACAGGAAATAAAAAGGGGAAATGAAACTGTTGAAATA 1813
Db 2035 TCTGTATATTTCCCGCGCTCGGCTCGCTCTTCTCAACGCTGTGATGCTGGCGACGCC 2094
Qy 1814 TATCAGTAAATGATCTGAGGCTGAAGGATTTGGGAAACATTTCTCAGACATTTGAAAAATATCAC 1873
Db 2095 GGCGCGATTTCCGGGATGCCAGAAAGTGTGTTCTGTCTCGCGCCGCCCATCGCTGATGA 2154
Qy 1874 TTTTACTCAAGGTCGACAAAGTCTCGCTGACTGCGGAAATGGAACCAACG----- 1923
Db 2155 AATCTCTATGCGCGCAACTGTGTGCGCTGCGAGAAATCTTTAACTGCGCGCGCGCGCA 2214
Qy 1924 -----ATCTGCGGCTCCAAACATGACCATCGGCGCTGTGAAACAGCAGAGAGTGAATTCACAGG 1978
Db 2215 GCGGATTTGCGCTCTGCGCTTCGCGAGGAGTCCGTACCGAAAGTGGATAAAAATTTTGG 2274
Qy 1979 CAGCTACATCAGACCGCTAAACAGCCACATCAAAATGAGATCA--AAGAGTCAACCACTGAT 2036
Db 2275 CCGCGCAACGCTTTTGTAAACCGAAGCCAAACGTCAGTCAAGCCAGGCTCTCGACGCGC 2334
Qy 2037 GGGACACAAAACACCATC-----AACAAAGAGACCCAGCCCACTTTGGCT 2082
Db 2335 GGTATCGATATGCGACCGCGCGCTGTGAAGTACTGATCGCAGACAGCGCGCGCAAC 2394
Qy 2083 TCACCGTCAATTTGGAAGTTTTCAGAGTCCACCACTGTCTTCAACGCGGCT----- 2130
Db 2395 ACCGATTTCTGCTCTGACCTGCTCTCCAGGCTGAGCACGCGCGGATTTCCAGGT 2454
Qy 2131 -----AGTCTTCTATAGACAGAAATGGGAAGAGTTCCTGAAGACCAT 2173
Db 2455 GATCCTGCTGACGCTGATGCTGACATTTGCGCAAGTGGCGAGGCGGTAGAACGTCA 2514
Qy 2174 GTGGCTGCTGGTCAAGTGTATGACATTTGGTGTGATGACTGGAAAGCTA---CCAGGGTC 2231
Db 2515 ACTGGCGGAATCGCGCGCGGACACCGCGCGGAGGCTGAGCGCGGATCTCTGAT 2574
Qy 2232 GGCAATCAACATCTTCACTCGCTGCGCACACAGAAAGAGTGTGAGTGAGTGAACAAAGTCT 2291
Db 2575 TGTGACCAAGATTTAGCGCAGTGGCTCGCCATCTCTAATCAGTATGGCGCGGAACACTT 2634
Qy 2292 CTGGAATCTCAGGTGAAAAAGGAGATAGAGCCCTCTCTGGACAA----- 2335
Db 2635 AATCATCCAGACGCGCAATCGCGCGATTTGGTGGATGCGATTACCAGCGCAGGCTCGGT 2694
Qy 2336 -----ANTGTTATACGAGCTTTCAGGCTTCAAGTCTAATAGTACTCCAGGTCTTAAAGG 2385
Db 2695 ATTTCTCGCGAGTGTGTCGCGGAAATCCGCGGTGATTACGCTTCCGGAACCAACCATGT 2754
Qy 2386 TGATCGGGGGATCTCTGTTTAACTCGAGTTCAGAGTTCAGAGGATTTCCAGGACCAATGGGAAGA 2445
Db 2755 TTTACGACCTATGGCTATCTGCTACTGTTCCAGGCTTTGGGTTAGGGATTTCCAGAA 2814
Qy 2446 CCGGAAACCGAGGACTTAATGGACAAAAGG-----CCAGAAAGGA 2486
Db 2815 ACGGATACCGCTTCAGGAACCTGTGAAAGCGGGCTTTTCCGCTCTGCGCATCAACCATGA 2874
Qy 2487 GAAAAAGGAGTGAAGCATGCAAGACATCTAATACAGTCCGACTGGTGGGTGGGAGC 2546
Db 2875 AACATTTGGCGCGCAGAACGTCTGACCGCCCATTAATAATGCGCTGACCCCTGCGCGTAAA 2934
Qy 2547 GGCCCTCAGAAAGCAGAGTG-----GAGATTTTTCAGAAAGCCAGTGGGGTACGGTG 2600
Db 2935 CGCCCTCAAGAGCAAGCATGAGCATGGAACACTCTCAGCGTCCGCGGATCAATTCG 2994
Qy 2601 TGTGACAGCCCTGGGAACTGCGTGGAGGACTG---GTGCTGTGCAAGGAGCTTGGGATACAA 2659
Db 2995 TTAACCTGAGGATCGGCGCCAGCCAGGATGGTATAGTATAGNATAGATTTTATTAGA 3054

Qy 2660 AGGTGTTCAAGTGTGCATAGCGAGCTTATTTTGGAAAGGTACGGTCCAATATGCT 2719
Db 3055 GCATAGGAGGGAGTCAAGAGGGGAGTAGAGGCAAGAGAGGAGAGGGGAGAG 3114
Qy 2720 GAATGAAGTATTTGTTTCGGGAAAGAGTC--ATCCATTGAAGAGTGCAGAAATAGACAG 2777
Db 3115 AAGAGAGTAGAGGTGTAGAGGCTGGCTTGGAGCACGTGGAGAGAGCAGGGAAGGAA 3174
Qy 2778 TGGGGTGTAGAGGCTGTTCGAC-----GACGAAGATGCTGGGGTCACT 2823
Db 3175 TGGGAGAGAGGAGNAGAAATTCGAAACAGAGAAAGCAAGAGAGGATCAAGAGAGCAAG 3234
Qy 2824 TTGCACCTACATATGATCATATTTTCATTCACATTTTAAACTGTTATTAAGTATTT 2883
Db 3235 AGGGGGCAAGAGCCGCTTTTATAGTGTAGTCAAGCACACCTGGTGTGTCCAGGTAACT 3294
Qy 2884 TTTTTCCTTTCTTCACTAAAA-----TCAGCTTAATTAATTAAGAAA 2929
Db 3295 GTGGGGGGGNTTAGACACAAAGCTAAACATCCACCACCTAGTTNATCTGCCAGACT 3354
Qy 2930 CTAAAGATTTT-----ATCCACAGAAAGGAATATTTAAATCA 2969
Db 3355 CTCAAGGTCCTCGCATACCAGAGCTGGGGAAATCCAGAGGGAGACATGTGAAGGCTCC 3414
Qy 2970 CTGGATMAACATATAAATAGCTTCATATTTGCTTCAATACCAAGAACCAATTT----- 3022
Db 3415 CTGGGACCTGCTCAGGACNGAGCAATGTGTGCCAGCNGAGTTCTCTAGGTCAAGC 3474
Qy 3023 -CAACTTCTCAGGTTTTAGTGGCTCGTGCAGAAATGATCCCTCAGGATA----- 3074
Db 3475 CTAAATGGCCAGATCTACACTTGGTGGCAGGAGTTTCAGGCTTTCTGGGAAGCAAAAC 3534
Qy 3075 -----TAGTAGTTTCGCTTTTGGCATAGGGAGGGGAAATGATGCTTATG 3119
Db 3535 TGGCAGAGAAACAGAGCAGGATCTTGATTTGGAAAGGAAGTCTAGGGCCAGAGACTGA 3594
Qy 3120 CAATACTCTTGTAGTCTTGAAATAGTGAACGATAGTAGTAGCAACATGCCCTTACAGGAG 3179
Db 3595 ACCTGGGGCTGCTGCTGTTCACCTGTNCTCCNGTGGTTTTATCCTCCAGTCTCTCAGC 3654
Qy 3180 AGAAAAGCACCGTGCATCGGATTTGTGGAGTAAGTGTGTAGCAGTCTGCTTATTAG 3239
Db 3555 CCCCTAGACCCAGCCAGCCATGACGCCGAGTAGCAAAAGTAAGAGGCCACAGGTCAG 3714
Qy 3240 GAAGGCAA---CAGACGGGTCTGACATGGAATGGACGAACCACTGAATTCGCGATTGCAG 3296
Db 3715 ACTGGTAGGCGACAGGTCCTGAGGCTTGGGGCAGGTGCTAGTTGGATCGGTTTCCAG 3774
Qy 3297 AGATATTGTATTTA-----AGTGCCCTAGCTCGATA 3326
Db 3775 AGGCAATATATATATAAGGCTGCTGTTTCCCGATGGTGCAACACCCAGAGGCTTAGC 3834
Qy 3327 CAGCAACGCCATTTGACCATTCACCATTTGGTGTGCACCTCCAGGTTCAAGCTGCGG 3386
Db 3835 CAGACCAACATTCCTGCGCAAAAGCCAGCTCTCCATTTATAGCACCTTGGAAAGTACGCC 3894
Qy 3387 CAAGCACTCAGGGCGCAGGGCTGTAAAGGAAGCGGAACACGTA----- 3431
Db 3895 AAGGAGCTGAATGCAAGGCTTGAAGAGCAGGGGGAATCAAGGGCCACCTATGGC 3954
Qy 3432 ----- 3431
Db 3955 GGCAGAGAACTGAGCCCTCAGGATGAGTGTGTCTTCTCCAGGTCAGTGGGCTACAGC 4014
Qy 3432 -----GAAAGCCAGTCCGAGAAACGGTGTGTGACCCCGGATGAATGTAGCTAC 3480
Db 4015 CTAAGAGGCCCTGATTTGAGGGGCAATGTCCTCAGCCAGAGCAATGGCAGCTCCAGA 4074
Qy 3481 TGGGCTATCTGCACAAGGGAAACGCAAGCGCAAGAGAGAAAGCAGG----- 3526
Db 4075 AGGGGTGCTGGGGAAGTTTTAGTGACCAATAAATAAAGCAAGGTTGCAATCACTTA 4134
Qy 3527 -----TAGCTTGAGTGGGCTT 3543

Db 4135 GACTCAGCATAAAAATTTATTTTGGTTTTNTAAATATATCGTTTGTATTTGTGCTT 4194
Qy 3544 ACATGGCGATAGCTAGACTTGGGCGGTTT----- 3571
Db 4195 GCTTTGCTTTGCTTTGTTTGTCTCGNGTAGACACTGGAGAGGCTTGGCGAGCTGCACCA 4254
Qy 3572 -----TATGGA 3577
Db 4255 GATAGCCAAGTGTACCGCGTATGGCTGCCAGTCTCTCTGCCCTCTTCTTAGCTAGGCA 4314
Qy 3578 CAGCAAGGAACCGAATTCAGCTGGGGCGCTCTGGTAAGGTTGGGAAGCCCTGCA 3637
Db 4315 GCTGCCCCAGCACAGATCGCGGAGGGGCACTCCCTTGGCCCCCAGTGGCTTACCTTGGG 4374
Qy 3638 AAGTAAA-----CTGGATGGCTTTCTTTCGCCGCCAAG 3668
Db 4375 ACCCAAGCTCCGCCCTACTACACTCTATTTGGGCTTGGGGGGGCTTTTAGGCTACCT 4434
Qy 3669 GATCTGATGGCGCAGG-----GATCAAGATCTCATCAAGAGACAGGATG 3713
Db 4435 GGATAAATAGCCAGAGCTTGGGCTGCAAGTCCGAGGCGGATCTCATCAAGAGACAGGATG 4494
Qy 3714 AGGATCGTTTGCATGATGAAAGATGATGATGCAAGCAGGTTCTCCGGCCGCTTGGGT 3773
Db 4495 AGGATCGTTTGCATGATGAAAGATGATGATGCAAGCAGGTTCTCCGGCCGCTTGGGT 4554
Qy 3774 GGAGAGGCTATTGGCTTATGACTGGGCAACACACAAATCGGCTGCTCTGATGCGCGCT 3833
Db 4555 GGAGAGGCTATTGGCTTATGACTGGGCAACACACAAATCGGCTGCTCTGATGCGCGCT 4614
Qy 3834 GTTCCGGCTGTGACGCGAGGGGCGCCGCTTTTGTCAAGACCGACCTGTCCGGTGC 3893
Db 4615 GTTCCGGCTGTGACGCGAGGGGCGCCGCTTTTGTCAAGACCGACCTGTCCGGTGC 4674
Qy 3894 CTTGAATGAATGACGAGACGAGCGCGCTATCTGGCTGCGCAGCAGCGGGGCTTCC 3953
Db 4675 CTTGAATGAATGACGAGACGAGCGCGCTATCTGGCTGCGCAGCAGCGGGGCTTCC 4734
Qy 3954 TTGGCAGCTGTGCTCCACGTTGTCAAGCGGGAAGGAGCTGGCTGCTATTGGGGCA 4013
Db 4735 TTGGCAGCTGTGCTCCACGTTGTCAAGCGGGAAGGAGCTGGCTGCTATTGGGGCA 4794
Qy 4014 AGTCCGGGCGAGGATCTCTGTCTCATCTCACCTTGTCTCTCGAGAAAGTATCCATCAT 4073
Db 4795 AGTCCGGGCGAGGATCTCTGTCTCATCTCACCTTGTCTCTCGAGAAAGTATCCATCAT 4854
Qy 4074 GGCTGATGCAATGCGGGGCTGCATAGCTTGTATCCGGCTACCTGCCCATTTGAGACCA 4133
Db 4855 GGCTGATGCAATGCGGGGCTGCATAGCTTGTATCCGGCTACCTGCCCATTTGAGACCA 4914
Qy 4134 AGCGAAACATCGCATCGAGCGAGCAGCTACTCGGATGGAAGCCGCTTGTTCGATCAGGA 4193
Db 4915 AGCGAAACATCGCATCGAGCGAGCAGCTACTCGGATGGAAGCCGCTTGTTCGATCAGGA 4974
Qy 4194 TGATCTGGAAGAGAGCATCAGGGGCTCGGCGCAGCCGAACTGTTCGCCAGGCTCAAGGC 4253
Db 4975 TGATCTGGAAGAGAGCATCAGGGGCTCGGCGCAGCCGAACTGTTCGCCAGGCTCAAGGC 5034
Qy 4254 GCGCATGCCGAGCGCGAGGATCTCGTGTGAACCCATGGCGATCGCTGTTCGCCAATAT 4313
Db 5035 GCGCATGCCGAGCGCGAGGATCTCGTGTGAACCCATGGCGATCGCTGTTCGCCAATAT 5094
Qy 4314 CATGGTGAAGTGGCGCTTTTCTGGATTCATCGATTCATCGATTCATCGATTCATCGATTC 4373
Db 5095 CATGGTGAAGTGGCGCTTTTCTGGATTCATCGATTCATCGATTCATCGATTCATCGATTC 5154
Qy 4374 CCGCTATCAGACATAGCTTGGCTACCGTGTATTTGCTGAAGAGCTTGGCGCGCAATG 4433
Db 5155 CCGCTATCAGACATAGCTTGGCTACCGTGTATTTGCTGAAGAGCTTGGCGCGCAATG 5214
Qy 4434 GGCTGACCGCTTCTCTGTGCTTTACGATTCGCGTATCGCGCTCCCGATTCGAGCGCATCGCTT 4493

QY 601 CCGGGATTGGAGACCCCTGCCAGGACCCAGGACCCACCCAGGAGGTAAGCTGGCC 660
DB 775 CCGGGATTGGAGACCCCTGCCAGGACCCAGGACCCACCCAGGAGGTAAGCTGGCC 834
QY 661 AGCAACTTATCTGTCTGTCCGATTTCTAGTGTCTATGTTGATGTTATGCGCTCGG 720
DB 835 AGCAACTTATCTGTCTGTCCGATTTCTAGTGTCTATGTTGATGTTATGCGCTCGG 894
QY 721 TCTGTACTAGTTAGTAACTAGTCTGTATCTGTGGCGGACCCGTGTGGAACCTGACAGTT 780
DB 895 TCTGTACTAGTTAGTAACTAGTCTGTATCTGTGGCGGACCCGTGTGGAACCTGACAGTT 954
QY 781 CTGACACCCCGCCGCAACCTCGGAGAGCTCCAGGACCTTGGGGCCCGTTTGTGG 840
DB 955 CTGACACCCCGCCGCAACCTCGGAGAGCTCCAGGACCTTGGGGCCCGTTTGTGG 1014
QY 841 CCGGACCTTGAGGAAGGAGTGCATGTGAATCCGACCCCGTCAGGATATGTGTTCTGTT 900
DB 1015 CCGGACCTTGAGGAAGGAGTGCATGTGAATCCGACCCCGTCAGGATATGTGTTCTGTT 1074
QY 901 AGGAGACGAGAACCTTAAACAGTTCCTCCGCTCCGTCGAAATTTTGTCTTTCGGTTTGGAA 960
DB 1075 AGGAGACGAGAACCTTAAACAGTTCCTCCGCTCCGTCGAAATTTTGTCTTTCGGTTTGGAA 1134
QY 961 CCGAAGCCCGGCTTGTCTGCTCGAGGCAAGCTTGGGCTGCAGTCTGACCTAGAGGA 1020
DB 1135 CCGAAGCCCGGCTTGTCTGCTCGAGGCAAGCTTGGGCTGCAGTCTGACCTAGAGGA 1162
QY 1021 TCAATTCGGCAGGAGTAATCGGTCGCTCGCTCTTTAGGACATATGAAGTATGGCAGGT 1080
DB 1163 TCAATTCGGCAGGAGTAATCGGTCGCTCGCTCTTTAGGACATATGAAGTATGGCAGGT 1202
QY 1081 GGGATGACTTTTCGATCAGCAAGAGGACATGACAGCTGTACAGAGTCTGTGAAGTTTCG 1140
DB 1203 TTCTGTATTTGTCGAAATTAGGSCCAGACTGTTACC-----ACTCCCTTAAGTTTG 1255
QY 1141 ATGCTCGCTCAGTGACAGCTTTGCTTCTCCCATCTTAAATAGGCCCAACTCTTCAAG 1200
DB 1256 ACCTTAGGTCACTGGAAGAGATGTCGAGCGGATCGCTCAACACCACTCGGTAGATGTCAG 1315
QY 1201 AGAGGATCAAGTCTTATAAAGTCACTGATCACTCCCTTTATCTCATTTGTTGTAGTTC 1260
DB 1316 AAGAGACGTTG-----GGTTACCTTCTGCTGCGAGAATGG 1351
QY 1261 TCGTGCCCATCATTTGGCAGTATGAGGAGTCTCAGCTCTCGAATGGGAAACGAAGATTGCA 1320
DB 1352 CCAACCTTTAAGCTCGGATGCGCGGAGAGCGGACCTTTAACCGAGACTCATCA----- 1406
QY 1321 CGTTGGCTCAGTTAATGCAGATATATCTCAAGTCCGGAAGGCAAGAAATGGCAGTG 1380
DB 1407 -----CCAGGT 1413
QY 1381 AAGATGAATCAGATTTCCAGNAGCTGTGATGGACGATGAGCAACATGGAAGCAGAA 1440
DB 1414 TAAGATCAAGGTCTTTTCACTGGCCCGCATGAGACACCCAGACAGGTCCCTCATCGT 1473
QY 1441 TCCAGTATCTTTCAGATATGAAGCAATCTCCTAGATGCTAAGAAATTTCCAAATTTCA 1500
DB 1474 GACCTGGGAAGCCTTGGCTTTTGACCCCTTCCCTGGGTCAAGCCCTTGTACACCCTAA 1533
QY 1501 GCATAACAACTGATCAAGAATTTAATGATGTTCTTTTCCAGCTAAATCTTACTTCCCT 1560
DB 1534 GCCTCGGCTCCTCTCTCTCCATCGGCCCGCTCTCTCCCTTGAACCTCCTCGTTGAC 1593
QY 1561 CCATCCAGAACATGAGATATCATAGGGATATCTCAAGTCAATTAGTATGCTGACACA 1620
DB 1594 CCGGCTCGATCTCCCTTTAT-----CCAGCCCTCACTCTCTCTAGG 1638
QY 1621 CCACAGTACTTGAATTTGAGTTTCAAGTATTGAACACTGAATGGCAGAGTCCAAGAGATG 1680
DB 1639 CCGCGGAATTCGTT-----AACT 1656
QY 1681 CATTTAAACAAGAGAGAGATGCGTAAATTTAGAGGAGCGTATATACAAATGCATCAGCAG 1740

DB 1657 CGAGGATCCGCTGTGGAAATGTGTCTAGTTAGGCTGTGGAAAGTCCCGAGGCTCCCCAG 1716
QY 1741 AAATTAAGTCTCTAGATGAAATAAAGTATATTTGGAAACAGAAATAAAGGGGAAATGA 1800
DB 1717 -----CAGGCAGAAGTATGCAAGCATGCATCTCAATTAGTCAGCA 1757
QY 1801 AACTGTTGAATATATCACTAATGATCTGAGGCTGAAGGATTTGGAAACATTTCTCAGACAT 1860
DB 1758 ACCAGGTGTGAAAGTCTCCAGGCTCCCGAGGAGCAGCAAGTATGCAAA---GCATGCAT 1814
QY 1861 TGAATAATATCATTCTACTCAAGGTGCCAAGAGTGTCTCGTGAATCTGGGAAATGGACCA 1920
DB 1815 CTCAATTAGTCAGCAACCATAGTCCCGCCCTTAATCTCCGCCCATCCCGCCCTTAATCTCGG 1874
QY 1921 ACATCTGGGCTCCAAACATGACCATCGGGCTGTGAACAGCAGAGAGGTGAATTTACAGGCA 1980
DB 1875 CCGAGTTCGGCCCATTTCTCCGCCCATGCTGACTAATTTTTTTTATTTATGACAGAGCC 1934
QY 1981 CCTACATCACAGCCGTAAACAGCCATCAAAATGAGATCAAAAGATCACCACTGCAATGGGA 2040
DB 1935 GAGCGCGCTCGGCTCTGAGCTATTCAGAAAGTAGTGAGGAGCT-----TTTTTGG 1987
QY 2041 CAAAAACACCATCAACAGAGAGACCCAGCCCACTTTGGCTTCACTGCTCAATTGGAAATG 2100
DB 1988 AGGCTTAGGCTTTTGCAAAAAGCTTGGGCTGCAAGCTTGGTACCGAGCTCGATCGATAT 2047
QY 2101 TTTCAGAGTCCACCACTGTCTTTCAGGSCCAGTGTCTTATAGACAGGAATGGAAGGAGG 2160
DB 2048 CTGGCGCGCTGACGAGTGTCTGGGCCCTCG-----AAATGGCGHACCT 2096
QY 2161 TCCTGAAGACCATGTGGCTGTGGGTCAAGTGTAAATGACATTTGATGATGATGGAAG 2220
DB 2097 TCCTGGACACCATGCGGTGGGCCAGGAGTACGCCGACGTACGCGATGA----- 2148
QY 2221 CTACAGGCTCGGCATCAAACTCTTCACTCCCTGCGCACACAGAGAGTGAAGTGTG 2280
DB 2149 -----GATCAATAAAGGGGCGCTGAGGA 2172
QY 2281 ACCAAGTCTCTGGACTCCAGGTGAAGAGATAGAGGCCCTCTCTGACAAAAATGG 2340
DB 2173 CCGGAGCGGCCAGAACCCGCTGACGACCCCGAGCGTCCCTGCTGCGCTCTCCCGG 2232
QY 2341 TATACAGGCTTTCCAGGCTTAATAGGTACTTCCAGGTCTTAAAGTGTATCGGGGGATCT 2400
DB 2233 GCTGCTGCCGAATCGCCCCAACGCATCTTGTGTTGTCACATCGAAGAACCGGCGG 2292
QY 2401 CTGTTTACCTGGAGTTCGAGGATTTCCAGGACCAATGGGGAGAACCGGGAGCCAGGAC 2460
DB 2293 GACCGTGACCGACAGTCCCGTAACTCCGTTAACCGTTGAAATTCAGTGGATCC----- 2346
QY 2461 TTAATGGCAAAAAGGCCAGAGGAGAAAAGGGAGTGGGAAGCATCAAGACAACTCA 2520
DB 2347 -----ACTA 2350
QY 2521 ATACAGTCCGACTGGTGGTGCGAGCGCCCTCACGAAGCAGAGTGGAGATTTTTCAG 2580
DB 2351 GTAACGCGCGCATGTGCTGG----- 2372
QY 2581 AAGGCCAGTGGGTTACGCTGTGTGACGACCCCTCGGAACTCGGTGGAGGACTGGTCTGCT 2640
DB 2373 -----AATTAATTCGCTGTCTGCGAGGCGCGGCTGTTGGGGTGAGTACTCCTCTCAAAA 2427
QY 2641 GCAGGAGCTTGGGATCAAAAGGTGTTCAAGTGTGCATAGCGAGCTTATTTTGGAAAAAG 2700
DB 2428 GCGGGCAGTACTTCTGCTGAAGATTTGTCAGTTTCCAAAAACAGAGGAGGATTTG----- 2481
QY 2701 GTACGGTCCAAATATGCTGAATGAAGTATTTTGTTCGGGAAAGAGTCACTCCATTTGAAG 2760
DB 2482 -----ATATTCACCTGCGCCGC 2498
QY 2761 AGTCAGAAATTTAGACGTGGGGTGTGAGAGCTTCTCGACGACGAAGAGATCTCGGGGTC 2820

Db 2499 GGTGATGCTTTAGGGTGGCGCGCTCCATCTGGTCAGAAAAAGACAAATCTTTTGTGTC 2558
Qy 2821 ACTTTGCACCTACATAAATGCAATCATATTTTCAATCACAATTTTAAACCTGATTAAGATG 2880
Db 2559 AAGCTTAGGTGTGGCAGGCTTGAGATCTGGCCATACACTTGAGTGACATGACATCCAC 2618
Qy 2881 ATTTTCTTCTTGTCTTCACTAAAATCAGCTTAATTAATATTTAAGAAACTAAGAAATTTT 2940
Db 2619 TTTGGCTTCTCTCC-----ACAGGT 2639
Qy 2941 ATCCACAGAAAGGAATATTTAAAAATCACTGGATAAACATATATAAATAGCTTCATATTT 3000
Db 2640 GTCCACTCCAGGTCCTCACTGAGGTGATCGAGCATGCATCTAGGGCGGCAATTCGCC 2699
Qy 3001 GCTTCAAAATACCAGAACCATTTCAACTTCTCTAGGTTTTAAAGTGGCTCGTGGCGAATTG 3060
Db 2700 CTTCTCCCTCCCGCCCTTAAC-----GTTACTGGCCGA 2734
Qy 3061 ATCCCTCAGGATATAGTAGTTTCGCTTTTGTATAGGAGGGGAAATGTAGTCTTATGC 3120
Db 2735 AGCGCTTGGAAATAGGCGGTGTGTGTCT-----ATATGTGATTTCCAC 2784
Qy 3121 AATACTCTTGTAGTCTTGCACATGGTAAACGATGAGTTAGCAACATGCTTACAAAGAGA 3180
Db 2785 CATATTGCCGCTTTTGGCAATGTGAGGGCCCGAAACCTGGCCCTGTCTTCTTGACGAG 2844
Qy 3181 GAAAAAGCACCGTGATCCCGATTGGTGGAAATAGGTGGTACGATCGTGCTTATTAGG 3240
Db 2845 CATTCCTAGGGTCTTTCCCTCTCGCCAAAGGA-----TGC 2882
Qy 3241 AAGGCAACAGACGGGTCTGACATGATTTGGACCAACCACTGAATTCGCCATTTGCAGAGAT 3300
Db 2883 AAGGTCTGTGAATGTCTGTAAGGAAGCAGTTCTCTGGAAGCTTCTTGAAGACAAACAA 2942
Qy 3301 ATTGTATTTAAGTGCCTAGCTGATPACAGCAAAACGCCATTTGACCAATTCACCAATTGGT 3360
Db 2943 GGTCTGAGCAACCTTTTGACGAGCAGCGAAACCCCACTGGCAGCAGGTGCCTCTCGC 3002
Qy 3361 GTGCACCTCCAGCTTACGCTGCCGCAAGCACTCAGGCGCCAGGGGTGCTAAAGGAAG 3420
Db 3003 GCCAAAGCCAGTGTATGAAGATACCTGCAAGGCGGCA----- 3045
Qy 3421 CGAAACAGGTAGAAACCAAGTCCGAGAAACGGTGTGACCCCGGATGAATGTTCAGCTAC 3480
Db 3046 ---ACCCAGTGCACGTTGTGAGTTGTAGTTGTGGAAGAGTCAATGCTCTCTC 3102
Qy 3481 TGGGCTATCTGCAAGGGAAGAAACCAAGCGCAAGAGAAAGACGAGTGTGCAAGTGGG 3540
Db 3103 AAGCGTAGTCAACAAAGGGGCTGAAGGATGCCAGAAAGGTACCCCAATTGTATGGGAATCTG 3162
Qy 3541 CTTACATGGCGATAGCTAGACTGGCGGTTTTATGACAGCAAGCAAGCGAACCGGAATTGCCA 3600
Db 3163 AT--CTGGGGCTCGGTGCATGCTTTACATGTTTGTAGTCGAGTTTAAAAAGCTCTA 3220
Qy 3601 GCTGGGGCGCCTCTGTGTAAGGTGTGGGAAGCCCTGCAAAAGTAAACCTGATGGCTTTCTTG 3660
Db 3221 GGCCCCCGAACCAAGGGAAGCTGTTTCTTTGAAAAACACGATGATAGCTTGCCAC 3280
Qy 3661 CGCCAAAGATCTGTATGGCGCAGGGGATCAAGATCTGATCAAGAGACAGGATAGGATCG 3720
Db 3281 AACCCCGGAT-----AATTCTCGCAGCCAAATATGGGATCG----- 3317
Qy 3721 TTTCCATGATTAACAAGATGGATTGCAAGGAGTTCTCCGGCGCTTGGGTGGAGAGG 3780
Db 3318 -----CAATTGAACAAGATGGATTGACGAGGTTCTCCGGCCGCTTGGGTGGAGAGG 3370
Qy 3781 CTTATCGGCTATGACTGGGCAACACAGACAAATCGGCTGCTCTGATGCGCGGTGTTCGG 3840
Db 3371 CTTATTCGGCTATGACTGGGCAACACAGACAAATCGGCTGCTCTGATGCGCGGTGTTCGG 3430
Qy 3841 CTGTACGGCAGGGGCGCCGGTCTCTTTTGTCAAGACCGACCTGTCCGGTGCCTGAAAT 3900
Db 3431 CTGTACGGCAGGGGCGCCGGTCTTTTGTCAAGACCGACCTGTCCGGTGCCTGAAAT 3490

Qy 3901 GAACTGCAGGACGAGGCGCGGCTATCGTGGCTGGCCAGCGGGCGGTTCTTTCGCA 3960
Db 3491 GAACTGCAGGACGAGGCGCGGCTATCGTGGCTGGCCAGCGGGCGTTCCTTTCGCA 3550
Qy 3961 GCTGTGCTCGAGCTTGTGATCTGAAGCGGGAAGGACCTGGCTGCTATTGGGCGAAGTGGCG 4020
Db 3551 GCTGTGCTCGAGCTTGTCACTGAAGCGGGAAGGACCTGGCTGCTATTGGGCGAAGTGGCG 3610
Qy 4021 GGGCAGGATCTTCCTGTCATCTCACCTTGTCTCCCGAGAAAGTATCCATCATGGCTGAT 4080
Db 3611 GGGCAGGATCTTCCTGTCATCTCACCTTGTCTCCCGAGAAAGTATCCATCATGGCTGAT 3670
Qy 4081 GCAATCGGGCGGCTGCAATCGCTTGTATCGGCTACCTGCCCATTCGACCAACCAAGCGAAA 4140
Db 3671 GCANTGGGGGCTGCAATCGCTTGTATCGGCTACCTGCCCATTCGACCAACCAAGCGAAA 3730
Qy 4141 CATCCATCGAGCAGCACTGATCTGATGGAAGCGGCTTGTGTCGATCAGATGATCTG 4200
Db 3731 CATCGCATCGAGCAGCACTGATCTGATGGAAGCGGCTTGTGTCGATCAGATGATCTG 3790
Qy 4201 GACGAAGCATCAGGGGCTCGCCAGCCGAACTGTCGACGCTCAAGCGCGCATG 4260
Db 3791 GACGAAGCATCAGGGGCTCGCCAGCCGAACTGTCGACGCTCAAGCGCGCATG 3850
Qy 4261 CCCGACGCGGAGGATCTCGTGTGACCATGGCGATGCTTGTGCCGAAATCATGTGTG 4320
Db 3851 CCCGACGCGGAGGATCTCGTGTGACCATGGCGATGCTTGTGCCGAAATCATGTGTG 3910
Qy 4321 GAAAAATGGCGCTTTCTGGAATCATCGACTGTGGCGGCTGGGTGTGGCGGACCGCTAT 4380
Db 3911 GAAAAATGGCGCTTTCTGGAATCATCGACTGTGGCGGCTGGGTGTGGCGGACCGCTAT 3970
Qy 4381 CAGGACATAGGCTTGGCTACCCGTCATATTTGCTGAAGCTTGGGGCGAATGGCTGAC 4440
Db 3971 CAGGACATAGGCTTGGCTACCCGTCATATTTGCTGAAGCTTGGCGGGAATGGCTGAC 4030
Qy 4441 CGCTTCTCTGTGCTTTACGGTATCGCGCTCCGATTCGACGCGCATCGCTTCTTATCGC 4500
Db 4031 CGCTTCTCTGTGCTTTACGGTATCGCGCTCCGATTCGACGCGCATCGCTTCTTATCGC 4090
Qy 4501 CTTCTTGAAGGATTC-----TTCTG 4520
Db 4091 CTTCTTGAAGGATTCGGTCGAGCGGATCTGATCAAGAGACAGGATGAGGATCGTTTCG 4150
Qy 4521 AGCGGATCTCGGGTTCGATATAAAGATTTTATTTAGTCTCCAGAAAGAGGGGG 4580
Db 4151 CGCGGACTCTGGGGTTCGATATAAAGATTTTATTTAGTCTCCAGAAAAAGGGGG 4210
Qy 4581 AATGAAAGACCCCACTGTAGGTTTGGCAAGCTAGCTTAAGTAAACGCCATTTTGAAGGC 4640
Db 4211 AATGAAAGACCCCACTGTAGGTTTGGCAAGCTAGCTTAAGTAAACGCCATTTTGAAGGC 4270
Qy 4641 ATGAAAAATAACATACTGAGAAATAGAGATTCAGATCAAGGTCAGGAAACAGATGGAAC 4700
Db 4271 ATGAAAAATAACATACTGAGAAATAGAGATTCAGATCAAGGTCAGGAAACAGATGGAAC 4330
Qy 4701 AGCTGAATATGGGCAACAGGATATCTGTGGTGAAGCATTTCTCTGCCCGGCTCAGGGCC 4760
Db 4331 AGCTGAATATGGGCAACAGGATATCTGTGGTGAAGCATTTCTCTGCCCGGCTCAGGGCC 4390
Qy 4761 AAGAACAGATGGAACAGCTGAATATGGCCAAAACAGGATATCTGTGGTGAAGCATTTCTCTG 4820
Db 4391 AAGAACAGATGGAACAGCTGAATATGGCCAAAACAGGATATCTGTGGTGAAGCATTTCTCTG 4450
Qy 4821 CCCCGGCTCAGGGCCAAAGAACAGATGTGTCCAGATGCGGTCCAGCCCTCAGCAGTTTCT 4880
Db 4451 CCCCGGCTCAGGGCCAAAGAACAGATGTGTCCAGATGCGGTCCAGCCCTCAGCAGTTTCT 4510
Qy 4881 AGAGAACCATCAGATGTTTCCAGGGTGGCCCAAGGACCTGAATGACCCCTGTGCTTAT 4940
Db 4511 AGAGAACCATCAGATGTTTCCAGGGTGGCCCAAGGACCTGAATGACCCCTGTGCTTAT 4570

QY 4941 TGAACCTACCAATCAGTTCGCTTCTCGCTTCTGTTGCGCGCTTCTGCTCCCGAGCTCA 5000
 Db |||||
 QY 4571 TGAACCTACCAATCAGTTCGCTTCTCGCTTCTGTTGCGCGCTTCTGCTCCCGAGCTCA 4630
 Db |||||
 QY 5001 ATAAAGAGAGCCCAACACCTCTCACTCGGGGCGCCAGTCTCCGATTGACTGAGTCGCCCG 5060
 Db |||||
 QY 4631 ATAAAGAGAGCCCAACACCTCTCACTCGGGGCGCCAGTCTCCGATTGACTGAGTCGCCCG 4690
 Db |||||
 QY 5061 GGTACCCGCTGATCAATAAACCCTCTTGCAGTTCATCCGACTTGTGGTCTCGCTGTC 5120
 Db |||||
 QY 4691 GGTACCCGCTGATCAATAAACCCTCTTGCAGTTCATCCGACTTGTGGTCTCGCTGTC 4750
 Db |||||
 QY 5121 CTTGGAGGGTCTCTCTGAGTGATGACTACCCGCTCAGCGGGGTCTTTTCAATTTGG 5177
 Db |||||
 QY 4751 CTTGGAGGGTCTCTCTGAGTGATGACTACCCGCTCAGCGGGGTCTTTTCAATTTGG 4807
 Db |||||

RESULT 7

AAV33629

ID AAV33629 standard; DNA; 5594 BP.

XX AC

XX AAV33629;

XX 29-DEC-1998 (first entry)

XX DT

XX DE

XX Gensa 981, a monomeric DNA sequence produced by the invention.

XX KW

XX Class IIS restriction endonuclease recognition site;

XX endogenous mouse promoter element; tissue-specific gene expression;

XX KW

XX hormone-specific gene expression; as;

XX KW

XX developmental-specific gene expression.

XX OS

XX Synthetic.

XX PN

XX WO9838326-A1.

XX PD

XX 03-SEP-1998.

XX PF

XX 28-FEB-1998; 98WO-05003918.

XX PR

XX 28-FEB-1997; 97US-0070910P.

XX PA

XX (NATU-) NATURE TECHNOLOGY CORP.

XX PI

XX Hodgson CP, Zink MA, Xu G;

XX PS

XX WPI, 1998-495399/42.

XX PT

XX Method for assembling gene or gene vector - comprises use of primers

XX PT

XX containing class IIS restriction endonuclease recognition sites.

XX PS

XX Example 2; Page 107-110; 141pp; English.

XX CC

XX The invention provides a novel method for directing self-assembly of a

XX CC

XX gene having three or more fragments in a directionally and spatially

XX CC

XX ordered fashion to produce a gene or a gene vector. The method involves

XX CC

XX usage of primers, containing class IIS restriction endonuclease

XX CC

XX recognition sites, for isolation of these fragments. As described in the

XX CC

XX disclosure, the method may also use a vector for the incorporation and

XX CC

XX screening of endogenous mouse promoter elements for the identification of

XX CC

XX cell specific promoters. In the example given, plasmids pBK-CMV

XX CC

XX (AAV33626), pVLM (AAV33623) and pVLOVNH-900 (AAV33621) were used as

XX CC

XX templates from which six fragments were amplified. Each of the fragments

XX CC

XX contained different regulatory sequences. The six PCR fragments were

XX CC

XX designed to self-assemble into a retro-vector using the method of the

XX CC

XX invention. The present sequence, designated as Gensa 981, represents the

XX CC

XX monomeric DNA sequence of the six ligated fragments. In general, the

XX CC

XX method is claimed to be useful for isolating and identifying regulatory

XX CC

XX sequences from a cell, including those for enhanced biological activity,

XX CC

XX or tissue-specific, hormone-specific or developmental-specific gene

XX CC

XX expression

XX SQ

XX Sequence 5594 BP; 1232 A; 1519 C; 1518 G; 1325 T; 0 U; 0 Other;

Query Match 31.5%; Score 1629.4; DB 2; Length 5594;
 Best Local Similarity 98.6%; Pred. No. 0;
 Matches 1654; Conservative 0; Mismatches 21; Indels 2; Gaps 1;
 QY 3371 AAGCTTACCGTCCGCGCAAGCACTCAGCGGCCAAGGGCTGCTAAAGGAAGCGGACACGT 3430
 Db |||||
 QY 3342 AAGCTTACCGTCCGCGCAAGCACTCAGCGGCCAAGGGCTGCTAAAGGAAGCGGACACGT 3401
 Db |||||
 QY 3431 AGAAGCCAGTCCCGCAGAAACGGTCTGACCCCGGATGAATGTCACTACTGGGCTATCT 3490
 Db |||||
 QY 3402 AGAAGCCAGTCCCGCAGAAACGGTCTGACCCCGGATGAATGTCACTACTGGGCTATCT 3461
 Db |||||
 QY 3491 GGAAGAAGGAAAAACGCAAGCGCAAGAAAGAGAGAGTGTGAGTGGGCTTACATGCG 3550
 Db |||||
 QY 3462 GGAAGAAGGAAAAACGCAAGCGCAAGAAAGAGAGAGTGTGAGTGGGCTTACATGCG 3521
 Db |||||
 QY 3551 GATAGCTAGACTGGGCGGTTTATGGACAGCAAGCGGAACCGGATTTGCCAGCTGGGGCGC 3610
 Db |||||
 QY 3522 GATAGCTAGACTGGGCGGTTTATGGACAGCAAGCGGAACCGGATTTGCCAGCTGGGGCGC 3581
 Db |||||
 QY 3611 CCTCTGGTAAGTTGGGGAAGCCCTGCAAAAGTAAACTGGATGGCTTTCTTCCGCCCAAGGA 3670
 Db |||||
 QY 3582 CCTCTGGTAAGTTGGGGAAGCCCTGCAAAAGTAAACTGGATGGCTTTCTTCCGCCCAAGGA 3641
 Db |||||
 QY 3671 TCTGATGCGCAGGGGATCAAGATCTGATCAAGAGACAGGATGAGGATCGTTTCCGATGA 3730
 Db |||||
 QY 3642 TCTGATGCGCAGGGGATCAAGATCTGATCAAGAGACAGGATGAGGATCGTTTCCGATGA 3701
 Db |||||
 QY 3731 TTGAACAAGATGGAATTCACGACAGGTTCTCCGGCCGCTTGGGTGAGAGGCTATTCGGCT 3790
 Db |||||
 QY 3702 TTGAACAAGATGGAATTCACGACAGGTTCTCCGGCCGCTTGGGTGAGAGGCTATTCGGCT 3761
 Db |||||
 QY 3791 ATGACTGGGCACAACAGACAACTGGCTGCTCTGATGCCCGCTGTTCCGGCTGTCCAGCGC 3850
 Db |||||
 QY 3762 ATGACTGGGCACAACAGACAACTGGCTGCTCTGATGCCCGCTGTTCCGGCTGTCCAGCGC 3821
 Db |||||
 QY 3851 AGGGCGCCCGGTTCTTTTGTCAAGACCGACCTGTCCGGTCCCTGAATGAATGCAGG 3910
 Db |||||
 QY 3822 AGGGCGCCCGGTTCTTTTGTCAAGACCGACCTGTCCGGTCCCTGAATGAATGCAGG 3881
 Db |||||
 QY 3911 ACAGGACAGCGCGGCTATCGTGGCTGGCCACAGCGGGGTTCTTCCGACCTGTGCTCG 3970
 Db |||||
 QY 3882 ACAGGACAGCGCGGCTATCGTGGCTGGCCACAGCGGGGTTCTTCCGACCTGTGCTCG 3941
 Db |||||
 QY 3971 ACCTTTGTCACCTGAAGCGGAAAGGAGTGGCTGCTATTGGGCGAAGTGCCGGCAGGATC 4030
 Db |||||
 QY 3942 ACCTTTGTCACCTGAAGCGGAAAGGAGTGGCTGCTATTGGGCGAAGTGCCGGGCGAGATC 4001
 Db |||||
 QY 4031 TCCTGTATCTCACCCTTGTCTCTCCCGAGAAAGTATCCATCATGCTGATGCAATCGCGC 4090
 Db |||||
 QY 4002 TCCTGTATCTCACCCTTGTCTCTCCCGAGAAAGTATCCATCATGCTGATGCAATCGCGC 4061
 Db |||||
 QY 4091 GGCTGCATACCTGTGATCCGGCTACCTGCCCATTCGACCAACAGCGGAAACATCGCATCG 4150
 Db |||||
 QY 4062 GGCTGCATACCTGTGATCCGGCTACCTGCCCATTCGACCAACAGCGGAAACATCGCATCG 4121
 Db |||||
 QY 4151 AGCCGACAGCTACCTCGGATGGAAGCGGCTCTTGTGATCAGGATGATCTGGACGAAGAGC 4210
 Db |||||
 QY 4122 AGCCGACAGCTACCTCGGATGGAAGCGGCTCTTGTGATCAGGATGATCTGGACGAAGAGC 4181
 Db |||||
 QY 4211 ATCAGGGGCTCGCGCCACAGCGAACTGTTTCCGACGCTCAAGCGCGCATGCCCGACGCGC 4270
 Db |||||
 QY 4182 ATCAGGGGCTCGCGCCACAGCGAACTGTTTCCGACGCTCAAGCGCGCATGCCCGACGCGC 4241
 Db |||||
 QY 4271 AGGATCTCGTGTGACCCATGCGCATGCTCTTCCGAAATATCATGTGGAATATGCGC 4330
 Db |||||
 QY 4242 AGGATCTCGTGTGACCCATGCGCATGCTCTTCCGAAATATCATGTGGAATATGCGC 4301
 Db |||||
 QY 4331 GCTTTTCTGGATTCATCGACTGTGGCGGCTGGGTGTGGCGGACCGCTATCAGGACATAG 4390
 Db |||||
 QY 4302 GCTTTTCTGGATTCATCGACTGTGGCGGCTGGGTGTGGCGGACCGCTATCAGGACATAG 4361
 Db |||||

QY 4391 CGTTGGCTACCGTGTATTTGCTGAAGAGCTTTGGCGGAAATGGGCTGACCGCTTCCTCG 4450
Db |||||
QY 4362 CGTTGGCTACCGTGTATTTGCTGAAGAGCTTTGGCGGAAATGGGCTGACCGCTTCCTCG 4421
Db |||||
QY 4451 TGTCTTACGGTATCGCCGCTCCGATTCGCAGCGCATCGCTTCTATCGCTTCTTGACG 4510
Db |||||
QY 4422 TGTCTTACGGTATCGCCGCTCCGATTCGCAGCGCATCGCTTCTATCGCTTCTTGACG 4481
Db |||||
QY 4511 AGTTCTTCTGAGCGGAGCTCTGGGTTGATATAAATAAAGATTTTA--TTTAGTCTTCA 4568
Db |||||
QY 4482 AGTTCTTCTGAGCGGAGCTCTGGGTTGATATAAATAAAGATTTTA--TTTAGTCTTCA 4541
Db |||||
QY 4569 GAAAGAGGGGGAATGAAAGAGCCCACTGTAGGTTTGGCAAGCTAGCTTAAGTAACGCC 4628
Db |||||
QY 4542 GAAAGAGGGGGAATGAAAGAGCCCACTGTAGGTTTGGCAAGCTAGCTTAAGTAACGCC 4601
Db |||||
QY 4629 ATTTTGAAGGATCGAAAAATACATACTGAGAAATAGAGAAGTTTCAGATCAAGGTCAAG 4688
Db |||||
QY 4602 ATTTTGAAGGATCGAAAAATACATACTGAGAAATAGAGAAGTTTCAGATCAAGGTCAAG 4661
Db |||||
QY 4689 AACAGATGGAACAGCTGATATGCGGCAACAGAGATATCTGTGTAAGCAGTTCTCTGCC 4748
Db |||||
QY 4662 AACAGATGGAACAGCTGATATGCGGCAACAGAGATATCTGTGTAAGCAGTTCTCTGCC 4721
Db |||||
QY 4749 CCGCTCAGGCGCAAGAACAGATGGAACAGCTGATATGCGGCAACAGAGATATCTGTG 4808
Db |||||
QY 4722 CCGCTCAGGCGCAAGAACAGATGGAACAGCTGATATGCGGCAACAGAGATATCTGTG 4781
Db |||||
QY 4809 AAGCAGTTCTCTCCCGCTCAGGCGCAAGAACAGATGTTCCCAAGATGCGGTCCAGGCC 4868
Db |||||
QY 4782 AAGCAGTTCTCTCCCGCTCAGGCGCAAGAACAGATGTTCCCAAGATGCGGTCCAGGCC 4841
Db |||||
QY 4869 TCAGCAGTTCTTAGAGAACCATCAGATGTTCCAGGTTGCCAGGACCTGAAATGACC 4928
Db |||||
QY 4842 TCAGCAGTTCTTAGAGAACCATCAGATGTTCCAGGTTGCCAGGACCTGAAATGACC 4901
Db |||||
QY 4929 CTGTGCTTATTGAACTAACCAATCAGTTCGCTTCTGCTTCTGTCGCGGCTTTCGC 4988
Db |||||
QY 4902 CTGTGCTTATTGAACTAACCAATCAGTTCGCTTCTGCTTCTGTCGCGGCTTTCGC 4961
Db |||||
QY 4989 TCCCGAGTCAATAAAGAGCCCAACCCCTCACTCGGGGCGCCAGTCTCTCGAT 5045
Db |||||
QY 4962 TCCCGAGTCAATAAAGAGCCCAACCCCTCACTCGGGGCGCCAGTCTCTCGAT 5018
Db |||||

RESULT 8

AAD28311
ID AAD28311 standard; DNA; 5130 BP.

AC AAD28311;

DT 22-APR-2002 (first entry)

DE LSRNL vector.

KW Regulatory element; vector; erythropoietin; growth hormone; insulin;
KW immunoglobulin; bone morphogenetic protein; interferon; interleukin;
KW superoxide dismutase; T-cell receptor; surface membrane protein;
KW viral antigen; transport protein; addressin; regulatory protein; MoMuLV;
KW moloney murine leukemia virus; chimeric; hepatitis B virus; ds.

OS Hepatitis B virus.
OS Moloney murine leukemia virus.
OS Unidentified.
OS Chimeric.

FT Key Location/Qualifiers
FT LTR 1..589

FT /*tag= a

FT /note= "MoMuLV 5' LTR"

FT misc_feature

FT /*tag= b

FT /note= "Retroviral packaging region"

FT misc_feature 1034..1714
FT /*tag= c
FT /note= "Hepatitis B surface antigen"
FT promoter 2279..2595
FT /*tag= d
FT /note= "RSV promoter"
FT misc_feature 2951..3745
FT /*tag= e
FT /note= "Neomycin phosphotransferase gene"
FT LTR 4537..5130
FT /*tag= f
FT /note= "MoMuLV 3' LTR"

W0200202783-A2.

10-JAN-2002.

29-JUN-2001; 2001WO-US020714.

03-JUL-2000; 2000US-0215851P.

(GALA-) GALA DESIGN INC.

Bleck GT;

WPI; 2002-154749/20.

Novel regulatory elements including nucleic acid encoding hybrid alpha-lactalbumin promoter or mutant RNA export element, for expressing one or more proteins e.g. antibodies, pharmaceutical proteins in host cells.

Example 1; Fig 12; 151pp; English.

The invention relates to novel regulatory elements and vectors for the expression of one or more proteins in a host cell. The invention further provides methods of indirectly detecting the expression of a protein of interest, comprising providing the host cell transfected with a vector encoding a polycistronic sequence comprising a signal protein and a desired protein operably linked by an internal ribosome entry site (IRES), and culturing the host cell under suitable conditions so that the signal protein and the desired protein is produced, where the presence of signal protein indicates the presence of desired protein. Regulatory elements and vectors of the invention are useful for the expression of proteins of interest in a host cell. They are useful for producing an immunoglobulin (Ig), preferably secretory Ig. They are useful in the expression of one or more proteins such as erythropoietin, growth hormone, insulin, immunoglobulin, protein C, cytokines and their receptors, hormones, Von Willebrand factor, lung surfactant, serum albumins, DNase, vascular endothelial growth factor, receptors for hormones or growth factors, rheumatoid factors, nerve growth factors, CD proteins, osteoinductive factors, immunotoxins, bone morphogenetic protein, interferons, colony stimulating factors, interleukins, viral superoxide dismutase, T-cell receptors, surface membrane proteins, antigens, transport proteins, addressins, regulatory proteins, antibodies, chimeric proteins and their fragments. The vectors are particularly useful for expressing G protein coupled receptors and other transmembrane proteins. The retroviral vectors are useful for expressing proteins in mammalian tissue culture host cells, including rat fibroblast cells, bovine kidney cells and human kidney cells. The present sequence is LSRNL vector used in the invention. The vector comprises the following elements: moloney murine leukemia virus (MoMuLV) 5' LTR, MoMuLV viral packaging signal, hepatitis B surface antigen, RSV promoter; neomycin phosphotransferase gene and 3' MoMuLV LTR

Sequence 5130 BP; 1187 A; 1335 C; 1332 G; 1276 T; 0 U; 0 Other;

Query Match 30.7%; Score 1588.6; DB 6; Length 5130;

Best Local Similarity 66.1%; Pred. No. 0;

Matches 3007; Conservative 0; Mismatches 754; Indels 786; Gaps 18;

Qy 1 TTTGAAAGACCCACCCGCTAGGTGCGCAAGCTAGCTTAAGTAACGCCACTTTGCAAGGCAT 60

Db 1 TTTGAAAGACCCACCCGCTAGGTGCGCAAGCTAGCTTAAGTAACGCCACTTTGCAAGGCAT 60

QY 4436 CTGACCGCTTCTCGTGTCTTACGGTATCGGCGCTCCCGATTCGAGCGCATCGCCTTCT 4495
DB |||||
DB 3660 CTGACCGCTTCTCGTGTCTTACGGTATCGGCGCTCCCGATTCGAGCGCATCGCCTTCT 3719
QY 4496 ATCGCTTCTTGACGAGTTCTTCTGAGCGGCACTCTCGGGTTCGATA 4542
DB |||||
DB 3720 ATCGCTTCTTGACGAGTTCTTCTGAGCGGCACTCTCGGGTTCGATA 3766

RESULT 9

ID ADM68972 standard; DNA; 5130 BP.

AC ADM68972;

DT 03-JUN-2004 (first entry)

XX LSRNL vector DNA.

XX Pharmaceutical; industry; diagnosis; LSRNL; ds.

XX Synthetic.

XX US2003224415-A1.

XX 04-DEC-2003.

XX 26-MAR-2003; 2003US-00397079.

XX 29-JUN-2001; 2001US-00897511.

XX 28-MAR-2002; 2002US-0368357P.

XX (GALA-) GALA DESIGN INC.

XX Brenel RD, Bleck GT, York D, Bakle K;

XX WPI; 2004-033957/03.

DR Host cell useful for producing proteins for pharmaceuticals, industrial
XX purposes, comprising genome having at least one integrated vector having
XX at least one exogenous gene and lacking selectable marker.

XX Example 1; SEQ ID NO 9; 78pp; English.

XX The invention relates to a host cell comprising a genome comprising at
CC least one integrated integrating vector, where the integrating vector
CC comprises at least one exogenous gene operably linked to a promoter, and
CC where the integrating vector lacks a gene encoding a selectable marker.
CC The invention also relates to a method of transfecting host cells,
CC involving providing a number of host cells comprising a genome and a
CC number of integrating vectors, where the integrating vectors comprise at
CC least one exogenous gene, and where the integrating vectors lack a gene
CC encoding a selectable marker, contacting the host cell with a number of
CC integrating vectors to generate transfected host cells comprising at
CC least one integrated copy of the integrating vector and selecting the
CC transfected host cells. The host cell is useful for producing a protein
CC of interest which involves providing a host cell, where the exogenous
CC gene encodes a protein of interest and culturing the host cell under
CC conditions such that the protein of interest is produced. The integrating
CC exogenous gene is stable in the absence of selection. The integrating
CC vector further comprises a secretion signal sequence operably linked to
CC the exogenous gene. The host cell is useful for the production of
CC proteins for pharmaceuticals and industrial, diagnostic and other
CC purposes, and in the production of multiple variants of proteins,
CC followed by analysis of the activity of the protein variants. This
CC sequence represents LSRNL vector DNA, used in the scope of the invention.

SQ Sequence 5130 BP; 1187 A; 1335 C; 1332 G; 1276 T; 0 U; 0 Other;

Query Match 30.7%; Score 1588.6; DB 12; Length 5130;

Best Local Similarity 66.1%; Pred. No. 0;

Matches 3007; Conservative 0; Mismatches 754; Indels 786; Gaps 18;

QY 1 TTGAAAGAGACCCACCCGTAGGTGGCAAGCTAGCTTAAGTAAACGCCACTTTGCAAGGCAT 60
DB |||||
DB 1 TTGAAAGAGACCCACCCGTAGGTGGCAAGCTAGCTTAAGTAAACGCCACTTTGCAAGGCAT 60
QY 61 GGAAAAATACATTAACCTAGGATAGAAAAGTTTCAGATCAAGGTCAAGGAACAAAGAAACAGC 120
DB |||||
DB 61 GGAAAAATACATTAACCTAGGATAGAAAAGTTTCAGATCAAGGTCAAGGAACAAAGAAACAGC 120
QY 121 TGAATACCAACAGGATATCTGTGTAAGCGGTTCTGCCCCCGGCTCAGGCGCCAGAAACA 180
DB |||||
DB 121 TGAATACCAACAGGATATCTGTGTAAGCGGTTCTGCCCCCGGCTCAGGCGCCAGAAACA 180
QY 181 GATGAGACAGCTGATGATGGGCCAAACAGGATATCTGTGTAAGCGGTTCTGCCCCCGG 240
DB |||||
DB 181 GATGAGACAGCTGATGATGGGCCAAACAGGATATCTGTGTAAGCGGTTCTGCCCCCGG 240
QY 241 CTGCGGGCCAGAACAGATGATGTTCCAGGCTGCCCCAGAGTCCAGGCTCAGAGTTCTAGTGAA 300
DB |||||
DB 241 CTGCGGGCCAGAACAGATGATGTTCCAGGCTGCCCCAGAGTCCAGGCTCAGAGTTCTAGTGAA 300
QY 301 TCATCAGATGTTCCAGGCTGCCCCAGAGGACTGAAAATGACCTGTACCTTATTTGAAC 360
DB |||||
DB 301 TCATCAGATGTTCCAGGCTGCCCCAGAGGACTGAAAATGACCTGTACCTTATTTGAAC 360
QY 361 TAACCAATCAGTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCT 420
DB |||||
DB 361 TAACCAATCAGTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCT 420
QY 421 AGAGCCCAACACCCCTCACTCGGGCGCCAGTCTTCGATAGACTGCGTCCCGGGTAC 480
DB |||||
DB 421 AGAGCCCAACACCCCTCACTCGGGCGCCAGTCTTCGATAGACTGCGTCCCGGGTAC 480
QY 481 CCGTATTTCCCAATAAAGCTCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTT 540
DB |||||
DB 481 CCGTATTTCCCAATAAAGCTCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTT 540
QY 541 GGAGGGTCTCTCTGAGTGAATTGATACCAACGAGGGGCTTTTCAATTTGGGGGCTCGT 600
DB |||||
DB 541 GGAGGGTCTCTCTGAGTGAATTGATACCAACGAGGGGCTTTTCAATTTGGGGGCTCGT 600
QY 601 CCGGGATTTGGAGACCCCTGCCCCAGGACCAACCCAGCCACCCGCGGAGTAAGCTGCCC 660
DB |||||
DB 601 CCGGGATTTGGAGACCCCTGCCCCAGGACCAACCCAGCCACCCGCGGAGTAAGCTGCCC 660
QY 661 AGCAACTTATCTGTGTCTGCTCCGATTTGCTAGTGTCTATGTTGATGTTATGCGGCTCG 720
DB |||||
DB 661 AGCAACTTATCTGTGTCTGCTCCGATTTGCTAGTGTCTATGTTGATGTTATGCGGCTCG 720
QY 721 TCTGTACTAGTTAGTAACTAGCTTCTGATCTGCGGACCCGCTGCTGGAACCTGACGAGTT 780
DB |||||
DB 721 TCTGTACTAGTTAGTAACTAGCTTCTGATCTGCTGCGGACCCGCTGCTGGAACCTGACGAGTT 780
QY 781 CTGAAACACCCGCGCGCAACCCCTGGGAGACGTCGCCAGGACTTTGGGGGCGCTTTTGTGG 840
DB |||||
DB 781 CTGAAACACCCGCGCGCAACCCCTGGGAGACGTCGCCAGGACTTTTGGGGGCGCTTTTGTGG 840
QY 841 CCGGACCTGAGGAGGAGGTCGATGCGAATCCGACCCCGCTCAGGATATGCTGCTTCTGCT 900
DB |||||
DB 841 CCGGACCTGAGGAGGAGGTCGATGCGAATCCGACCCCGCTCAGGATATGCTGCTTCTGCT 900
QY 901 AGGAGACGAGAACCTAAACAGATTTCCCGCTCTCGATTTTGTGCTTTTGGTTTGGAA 960
DB |||||
DB 901 AGGAGACGAGAACCTAAACAGATTTCCCGCTCTCGATTTTGTGCTTTTGGTTTGGAA 960
QY 961 CCGAAGCGCGGCTGCTGCTGTCGACCAAGCTTGGGCTGCGAGTCTGAGGTA 1020
DB |||||
DB 961 CCGAAGCGCGGCTGCTGCTGTCGACCAAGCTTGGGCTGCGAGTCTGAGGTA 1020
QY 1021 TCAATTCGGCACGAGTAAATCGGTGCTGCTTCTTAGGACATATGAGTATGCACAGT 1080
DB |||||
DB 1021 CC-----CTGACCAGAACATGGAGAACACACATCAGGATTCCT 1059

Db 2460 AAGCAAACAGACAGCTCTGACATGATGGAGCAACCACTGAAATTCGCGATTGACAGAT 2519
Qy 3301 -ATTGTAATTAAGTCCCTAGCTCGATACAGCAACGCCA--TTTGACCAATTCACACATT 3357
Db 2520 AATTGTAATTAAGTCCCTAGCTCGATACAGCAACGCCAATTTTGACCAATTCACACATT 2579
Qy 3358 GGTGTGCACT--CCAAAGCTTCACGCTCGCGCAAGCACTCAGGCGCGCAAGGCTGCTAAA 3415
Db 2580 GGTGTGCACTTCCAAAGCTTCACGCTCGCGCAAGCACTCAGGCGCGCAAGGCTGCTAAA 2639
Qy 3416 GGAAGCGAAACACGTAGAAAGCCAGTCCGCGAAGAAAGCGTGTGACCCCGGATGAATGTCA 3475
Db 2640 GGAAGCGAACAACGTAGAAAGCCAGTCCGCGAAGAAAGCGTGTGACCCCGGATGAATGTCA 2699
Qy 3476 GCTACTGGGCTATCTGGCAAGAGGAAACGCAAGCGCAAGAAAGACAGGTAGCTTGCA 3535
Db 2700 GCTACTGGGCTATCTGGCAAGAGGAAACGCAAGCGCAAGAAAGACAGGTAGCTTGCA 2759
Qy 3536 GTGGGCTTACATGGCGATAGCTAGACTGGGCGGTTTATGACACCAAGCGAAT 3595
Db 2760 GTGGGCTTACATGGCGATAGCTAGACTGGGCGGTTTATGACACCAAGCGAAT 2819
Qy 3596 TGCCAGCTGGGCGGCGCTCTGTAGAGTTGGAGCCCTGCAAAAGTAACTGGATGGCTT 3655
Db 2820 TGCCAGCTGGGCGGCGCTCTGTAGAGTTGGAGCCCTGCAAAAGTAACTGGATGGCTT 2879
Qy 3656 TCTTGCCGCCAAGGATCTGATGGCGCAGGGATCAAGATCTGATCAAGACAGAGATGAG 3715
Db 2880 TCTTGCCGCCAAGGATCTGATGGCGCAGGGATCAAGATCTGATCAAGACAGAGATGAG 2939
Qy 3716 GATGCTTTCGATGATGAACAAGATGATGACGACGAGTTCTCCGCGCGCTGGGTGG 3775
Db 2940 GATGCTTTCGATGATGAACAAGATGATGACGACGAGTTCTCCGCGCGCTGGGTGG 2999
Qy 3776 AGAGGCTATTGGCTATGACTGGGCACAAACAGACAACTGGCTGCTCATGCCGCGGT 3835
Db 3000 AGAGGCTATTGGCTATGACTGGGCACAAACAGACAACTGGCTGCTCATGCCGCGGT 3059
Qy 3836 TCCGCTGTACGCGCAGGGGCGCGGTTCTTTTGTCAAGACCGACCTGCCGCTGCC 3895
Db 3060 TCCGCTGTACGCGCAGGGGCGCGGTTCTTTTGTCAAGACCGACCTGCCGCTGCC 3119
Qy 3896 TGAATGAATGACGACAGGACGCGCGCTATCGTGGCTGGCACAAGCGCGGCTTCCTT 3955
Db 3120 TGAATGAATGACGACAGGACGCGCGCTATCGTGGCTGGCACAAGCGCGGCTTCCTT 3179
Qy 3956 GCGACGCTGCTCGACCTTGTCACTGAAGCGGAGGGAATGGCTGCTATGGCGGAAG 4015
Db 3180 GCGACGCTGCTCGACCTTGTCACTGAAGCGGAGGGAATGGCTGCTATGGCGGAAG 3239
Qy 4016 TGCCGGGCGCAGGATCTCCTGTCTATCTCACCTTGTCTCTCGCGAGAAAGTATCCATCATGG 4075
Db 3240 TGCCGGGCGCAGGATCTCCTGTCTATCTCACCTTGTCTCTCGCGAGAAAGTATCCATCATGG 3299
Qy 4076 CTGATGCAATGCGCGGCTGATACGCTTGATTCGCGCTACTGCGCATTCGACCAACCAAG 4135
Db 3300 CTGATGCAATGCGCGGCTGATACGCTTGATTCGCGCTACTGCGCATTCGACCAACCAAG 3359
Qy 4136 CGAATCATGCTCAGCGGACGACGTACTCGGATGGAAGCGGCTTGTGCTCATGAGATG 4195
Db 3360 CGAATCATGCTCAGCGGACGACGTACTCGGATGGAAGCGGCTTGTGCTCATGAGATG 3419
Qy 4196 ATCTGGACGAGAGATCAGGGGCTCGCGCGCAGCGCAAGCTGTTGCCAGGCTCAAGGGC 4255
Db 3420 ATCTGGACGAGAGATCAGGGGCTCGCGCGCAGCGCAAGCTGTTGCCAGGCTCAAGGGC 3479
Qy 4256 GCATGCCCGACGGGAGGATCTGCTGTGAACCCATGCGCATGCTGCTTGGCGAATATCA 4315
Db 3480 GCATGCCCGACGGGAGGATCTGCTGTGAACCCATGCGCATGCTGCTTGGCGAATATCA 3539
Qy 4316 TGGTGGAAATGCGCGCTTTCTGATTCATGCACTGTGGCGGCTGGGTGGCGGACC 4375
Db 3540 TGGTGGAAATGCGCGCTTTCTGATTCATGCACTGTGGCGGCTGGGTGGCGGACC 3599

Qy 4376 GCTATCAGGACATAGCGTTGGCTACCGGTGATATTCGTGAAGAGCTTGGCGCGAATGGG 4435
Db 3600 GCTATCAGGACATAGCGTTGGCTACCGGTGATATTCGTGAAGAGCTTGGCGCGAATGGG 3659
Qy 4436 CTGACCGCTTCTCGTCTTTACCGTATCGCGCTCCCGATTCGACGCGCATCGCTTCT 4495
Db 3660 CTGACCGCTTCTCGTCTTTACCGTATCGCGCTCCCGATTCGACGCGCATCGCTTCT 3719
Qy 4496 ATCCGCTTCTTGACGAGTTCTTCTGAGCGGACCTCTGGGGTTGATA 4542
Db 3720 ATCCGCTTCTTGACGAGTTCTTCTGAGCGGACCTCTGGGGTTGATA 3766

RESULT 10

ID ADW28107
ID ADW28107 standard; DNA; 5130 BP.
XX
AC ADW28107;
XX
DT 07-APR-2005 (first entry)
XX
DE Nucleotide sequence of LSNRL vector.
XX
KW recombinant protein; integrating vector; LSNRL vector; ss.
XX
OS Synthetic.
XX
FH Key
FT LTR Location/Qualifiers
FT 1..589
FT /*tag= a
FT /note= "MoMuSV 5'LTR"
FT 659..897
FT misc_feature /*tag= b
FT /note= "retroviral packaging region"
FT 1034..1714
FT /*tag= c
FT /product= "Hepatitis B surface antigen"
FT 2279..2595
FT /*tag= d
FT /note= "RSV promoter"
FT 2951..3745
FT CDS /*tag= e
FT /gene= "neomycin phosphotransferase gene"
FT 4537..5130
FT LTR /*tag= f
FT /note= "MoMuSV 3'LTR"

WO2005007803-A2.

27-JAN-2005.

27-MAR-2003; 2003WO-US009325.

28-MAR-2002; 2002US-0368357P.

26-MAR-2003; 52US-00368357.

(GALA-) GALA DESIGN INC.

Bremel Robert D, York D, Bleck GT, Bakle KA;

WPI; 2005-102089/11.

New host cell comprising integrated integrating vector, used in the
absence of selection and for producing high levels of proteins.

Example 1; SEQ ID NO 9; 148pp; English.

The invention relates to the production of proteins in host cells. The
host cells contain multiple integrated copies of an integrating vector
comprising an exogenous gene operably linked to a promoter, and where the
integrating vector lacks a gene encoding a selectable marker. The
invention is used to produce proteins of interest at high levels. The

CC present sequence represents a LSURL vector. This vector is used to
CC produce proteins in the method of the invention.

SQ Sequence 5130 BP; 1187 A; 1335 C; 1332 G; 1276 T; 0 U; 0 Other;

Query Match 30.7%; Score 1588.6; DB 14; Length 5130;

Best Local Similarity 66.1%; Pred. No. 0;

Matches 3007; Conservative 0; Mismatches 754; Indels 786; Gaps 18;

| | | | |
|----|-----|--|-----|
| Qy | 1 | TTTGAAGACCCACCCAGCTAGTGGTGAAGCTAGCTTAAGTAAAGCCACTTTGCAAGGCAT | 60 |
| Db | 1 | TTTGAAGACCCACCCAGCTAGTGGTGAAGCTAGCTTAAGTAAAGCCACTTTGCAAGGCAT | 60 |
| Qy | 61 | GGAAAAATACATAACTGAGATAGAAAAGTTTCAGATCAAGTTCAGGACCAAGAAACAGC | 120 |
| Db | 61 | GGAAAAATACATAACTGAGATAGAAAAGTTTCAGATCAAGTTCAGGACCAAGAAACAGC | 120 |
| Qy | 121 | TGAATACCAACACAGGATATCTGTGTAAAGCGTTCTCCCGCGCTCAGGGCCCAAGACA | 180 |
| Db | 121 | TGAATACCAACACAGGATATCTGTGTAAAGCGTTCTCCCGCGCTCAGGGCCCAAGACA | 180 |
| Qy | 181 | GATGAGACAGCTGAGTGAATGGCCAAAACAGGATATCTGTGTAAAGCGTTCTCCCGCGG | 240 |
| Db | 181 | GATGAGACAGCTGAGTGAATGGCCAAAACAGGATATCTGTGTAAAGCGTTCTCCCGCGG | 240 |
| Qy | 241 | CTCGGGCCCAAGACAGATGTCCTCCAGATCGGCTCCAGCCCTCAGCAGTTCTAGTGAA | 300 |
| Db | 241 | CTCGGGCCCAAGACAGATGTCCTCCAGATCGGCTCCAGCCCTCAGCAGTTCTAGTGAA | 300 |
| Qy | 301 | TCATCAGATGTTTCAGGGTGCCTCCAGGACCTGAAATGACCCCTGTACCTTATTTGAAC | 360 |
| Db | 301 | TCATCAGATGTTTCAGGGTGCCTCCAGGACCTGAAATGACCCCTGTACCTTATTTGAAC | 360 |
| Qy | 361 | TAAACCAATCAGTTCGCTTCCTGCTTCCTGCTTCGCGCTTCGCTCTCCGAGCTCAATAA | 420 |
| Db | 361 | TAAACCAATCAGTTCGCTTCCTGCTTCCTGCTTCGCGCTTCGCTCTCCGAGCTCAATAA | 420 |
| Qy | 421 | AGAGCCCAACACCCCTCACTCGGCGCGCCAGTCTTCGATAGACTGCGTCCCGGGTAC | 480 |
| Db | 421 | AGAGCCCAACACCCCTCACTCGGCGCGCCAGTCTTCGATAGACTGCGTCCCGGGTAC | 480 |
| Qy | 481 | CCGTATTCCTCAATAAGCCTCTTGCTTTGATCCGAATCGTGTCTCGCTTCCTCTTG | 540 |
| Db | 481 | CCGTATTCCTCAATAAGCCTCTTGCTTTGATCCGAATCGTGTCTCGCTTCCTCTTG | 540 |
| Qy | 541 | GGAGGGTCTCTCTGAGTGAATGACTACCCACGAGCGGGCTTTTCATTTCGGGGCTCGT | 600 |
| Db | 541 | GGAGGGTCTCTCTGAGTGAATGACTACCCACGAGCGGGCTTTTCATTTCGGGGCTCGT | 600 |
| Qy | 601 | CCGGATTTGGAGACCCCTGCCAGGACCAACGACCCACCCAGGGAGGTAACTGGCC | 660 |
| Db | 601 | CCGGATTTGGAGACCCCTGCCAGGACCAACGACCCACCCAGGGAGGTAACTGGCC | 660 |
| Qy | 661 | AGCAACTTATCTGTCTGTCCGATTGTCTAGTGTCTATGTTGATGTTATGCGCCTCGC | 720 |
| Db | 661 | AGCAACTTATCTGTCTGTCCGATTGTCTAGTGTCTATGTTGATGTTATGCGCCTCGC | 720 |
| Qy | 721 | TCTGTACTAGTAACTAGCTCTGATCTGCGGACCCCGTGGTGAACCTGACGAGTT | 780 |
| Db | 721 | TCTGTACTAGTAACTAGCTCTGATCTGCGGACCCCGTGGTGAACCTGACGAGTT | 780 |
| Qy | 781 | CTGAACACCCCGCCCAACCTTGGAGACTCCCAAGGACCTTTGGGGCCGTTTTGTGG | 840 |
| Db | 781 | CTGAACACCCCGCCCAACCTTGGAGACTCCCAAGGACCTTTGGGGCCGTTTTGTGG | 840 |
| Qy | 841 | CCCGACCTGAGAGAGGGAGTGCATGTGGAATCCGACCCCGTCAGGATATGTGTTCTGGT | 900 |
| Db | 841 | CCCGACCTGAGAGAGGGAGTGCATGTGGAATCCGACCCCGTCAGGATATGTGTTCTGGT | 900 |
| Qy | 901 | AGGAGACGAGAACCTTAAACAGTTTCCCGCTCCGCTCTGAATTTTTCGTTTCGTTTGAA | 960 |
| Db | 901 | AGGAGACGAGAACCTTAAACAGTTTCCCGCTCCGCTCTGAATTTTTCGTTTCGTTTGAA | 960 |

| | | | |
|----|------|--|------|
| Qy | 961 | CCGAAGCGCGCGCTTGTCTGTGACGACCAAGCTTGGGCTGCGAGTTCGACTTAGAGGA | 1020 |
| Db | 961 | CCGAAGCGCGCGCTTGTCTGTGACGACCAAGCTTGGGCTGCGAGTTCGACTTAGAGGA | 1020 |
| Qy | 1021 | TCAATTCGGCACGAGTAAATCGGTGCTCTTTAGGACATATGAAGTATGGACAGT | 1080 |
| Db | 1021 | CC-----CTGCACCGAACATGGAGAACACACATCAGGATTCCT | 1059 |
| Qy | 1081 | GGATGACTTTCCTGATCAGCAGAGGACACTGACAGCTGTACAGAGTCTGTGAAGTCG | 1140 |
| Db | 1060 | AGGACCCCTCTCGTGTACAGCGGGGTTTTCTTGTGACAAGAAATCTTCACAATACC | 1119 |
| Qy | 1141 | ATGCTCGCTCAGTGACAGCTTTCCTCCTCCCACTCTAAATAATGGGCAACTCTTCAAG | 1200 |
| Db | 1120 | ACAGAGTCTAGACTCGGTGGACTTCTCTCAATTTCTAGGGGAGGACCCACGTGTCC | 1179 |
| Qy | 1201 | AGAGGATGAAGTCTTATAAAGCTGCACTGATCAACCTTTTATCTCATCTGTGTGTAGTTC | 1260 |
| Db | 1180 | TGGCAAAAATTCGCGAGTCCCAACCTCAATCACTACCAACCTCTTGTCTCCCAATTTG | 1239 |
| Qy | 1261 | TGCTGCCCATCATTTGGCATAGTGGCAGCTCAGCTCTCTGAAATGGGAAACGAGAAATTGCA | 1320 |
| Db | 1240 | TCCTGGCTATC-----GCTGGATGTCTGCGGCGTTTTATC | 1276 |
| Qy | 1321 | CGGTTGGCTCAGTTAATGACAGATATATCTCCAAGTCCGGAAGCAAAAGGAAATGGCAGTG | 1380 |
| Db | 1277 | ATAITCTCTTCACTCTGCTGCTATGCTCATCTTCTTGTGTGTCTTCTGGACTACCA | 1336 |
| Qy | 1381 | AAGATGAAGTTCGAGATTCGAGAGAGCTGTGTAAGCAACGATGAGCAACATGAAAGCAGAA | 1440 |
| Db | 1337 | GGTATGTTGCGCGTTTGTCTCTACTTCCAGGAA--CATCAACTACAGCACGGGACCA | 1393 |
| Qy | 1441 | TCCAGTATCTTTCAGATAATGAAGCCAAATCTCTAGATGCTAAGAAATTCGAAATTTCA | 1500 |
| Db | 1394 | TGCAAGACCTGCGAGATCTCTGCTCAAGGAACCTCTATGTTTCCCTCTT-----GTTGCT | 1448 |
| Qy | 1501 | GCATAACAATGATCAAGATTTAATGATGTTCTTTTCCAGCTAAATTCCTTACTTTCCCT | 1560 |
| Db | 1449 | GTACAAAACCTTCGGACGGAAACCTGCACTTGATATCCCATCCCATCATCTCTGGGCTTCG | 1508 |
| Qy | 1561 | CCATCCAGGAACATGAGAAATATCATAGGGATATCTCCAAGTCAATTAAGTCTGAACA | 1620 |
| Db | 1509 | CAAGATTCCTATGGAGTGGGCTCAGTCCGTTCTCTCTCG----- | 1549 |
| Qy | 1621 | CCACAGTACTTGATTTGCGAGTTCAGTATTGAAACACATGAAATGGCAGAGTCCAGAGATG | 1680 |
| Db | 1550 | -CTCAGTTTACTAGTGCATTTGTTTCAGTGGTTTC---GTAGGGCTTTCCCGCACTGTTG | 1605 |
| Qy | 1681 | CATTAAACAACAAGAGGAGATCGTAAATTTAGAGGAGCGTATATACAATGCATCAGCAG | 1740 |
| Db | 1606 | GCITTCAGTTATG-GATGATGTGGTATTTGGGGGCCAAGTCTGTACACATCTTGAGTC | 1664 |
| Qy | 1741 | AAATTAAGTCTCTAGATGAAAAACAAGTATATTTGGAAACAGGAAATAAAAGGGGAAATGA | 1800 |
| Db | 1665 | CTTTTTCCTCTATTACCAATTTCTTTGTTCTTGGGTATACATTTAAACCCCTAATAA | 1724 |
| Qy | 1801 | AACGTGTAATATATCATTAATGATCTGAGGCTGAAGGATGGGAAACATTTCTCAGACAT | 1860 |
| Db | 1725 | AAC-----CAAAACGT | 1734 |
| Qy | 1861 | TGAAAAATATACATTTTACTCCAGGTGCCAGAAAGTCTCGCTGCTGGGAAATGGACCA | 1920 |
| Db | 1735 | TGGGGCTACTCCCTTAACTTCATGGGATATGTAATTTGATGTTGGGG----- | 1781 |
| Qy | 1921 | ACGATCTGGGCTCAACATGACCATCGGGGCTGTGAACACGACAGAGGTGAATTCACAGSCA | 1980 |
| Db | 1782 | ----- | 1781 |
| Qy | 1981 | CCTACATCAGCGGTAAAGCCACATCAAAATGAGATCAAAAGATCAACCTGATGAGGA | 2040 |
| Db | 1782 | -----TACTTTACCGCAAGAACATATTGTACTAAAAATCAAGCAATGTTTTCG- | 1829 |
| Qy | 2041 | CACAAAACACCATCAACAAGAGGAGCCCGCCACCTTTTGGCTTCAACGTCATTTGGAAGT | 2100 |


```
QY 4256 GCATGCCGACGGGAGGATCTGCTGCTGACCATGGCGATGCTGCTTGCAGATATCA 4315
DB 3480 GCATGCCGACGGGAGGATCTGCTGCTGACCATGGCGATGCTGCTTGCAGATATCA 3539
QY 4316 TGGTGAAATAGGCGCTTTTCTGGAATCATCGACTGTGGCGGCTGGGTGTGGGAC 4375
DB 3540 TGGTGAAATAGGCGCTTTTCTGGAATCATCGACTGTGGCGGCTGGGTGTGGGAC 3599
QY 4376 GCTATCAGACATAGCTTGGCTACCCCTGATATTTCTGAAGAGCTTGGCGGCAATGG 4435
DB 3600 GCTATCAGACATAGCTTGGCTACCCCTGATATTTCTGAAGAGCTTGGCGGCAATGG 3659
QY 4436 CTGACCGCTTCTCTGCTGCTTTACGGTATCGCGGCTCCCGATTCCGAGGCAATCGCTTCT 4495
DB 3660 CTGACCGCTTCTCTGCTGCTTTACGGTATCGCGGCTCCCGATTCCGAGGCAATCGCTTCT 3719
QY 4496 ATCGCTTCTTACGAGTCTTCTTACGAGCGGACTCTGGGGTTCGATA 4542
DB 3720 ATCGCTTCTTACGAGTCTTCTTACGAGCGGACTCTGGGGTTCGATA 3766

RESULT 11
AEB77874
ID AEB77874 standard; DNA; 5130 BP.
XX AC AEB77874;
XX
DT 20-OCT-2005 (first entry)
XX
DE Retroviral vector expressing hepatitis B surface antigen.
XX
KW Vector; gene therapy; cell transduction; transfection;
KW protein production; ds; surface antigen.
XX
OS Moloney murine sarcoma virus.
OS Respiratory syncytial virus.
OS Moloney murine leukemia virus.
OS Hepatitis B virus.
OS Synthetic.
OS Unidentified.
XX
FH Key Location/Qualifiers
FT LTR 1. 589
FT /tag= a
FT /note= "MoMuSV 5' LTR"
FT misc_feature 659. .897
FT /tag= b
FT /note= "Retroviral packaging region"
FT CDS 1034. .1714
FT /tag= c
FT /product= "hepatitis B surface antigen"
FT promoter 2279. .2595
FT /tag= d
FT /note= "RSV promoter"
FT misc_feature 2951. .3745
FT /tag= e
FT /note= "Neomycin phosphotransferase gene"
FT LTR 4537. .5130
FT /tag= f
FT /note= "MoMuLV 3' LTR"
XX
PN WO2005072129-A2.
XX
PD 11-AUG-2005.
XX
PF 14-JAN-2005; 2005WO-US001165.
XX
PR 16-JAN-2004; 2004US-00759315.
XX
PA (GALA-) GALA DESIGN INC.
XX
PI Black GT, Bremel RD, Miller LU;
```

```
XX WPI; 2005-555595/56.
DR Transducing host cells useful for producing essential proteins comprises
PT serially contacting host cells with integrating retroviral vectors to
PT provide host cells comprising multiple integrated retroviral vectors.
XX
PS Example 1; SEQ ID NO 9; 168pp; English.
XX
CC The invention relates to transducing host cells comprising providing at
CC least one host cell comprising a genome and a population of retroviral
CC vectors encoding a gene of interest, contacting the host cell with the
CC population of integrating vectors (under conditions such that the host
CC cells are transduced to produce transduced host cells) and repeating both
CC steps several times (e.g. at least 4-20) to provide host cells comprising
CC multiple integrated retroviral vectors (about 10-100 integrated
CC retroviral vectors). Also included is the host cell produced by the above
CC method. The genes of interest are arranged in a polyclonic sequence,
CC and comprise immunoglobulin heavy and light chains. The integrating
CC vector further comprises a secretion signal sequence operably linked to
CC the exogenous gene. The retroviral vector further encodes an amplifiable
CC marker selected from dihydrofolate reductase (DHFR) and glutamine
CC synthetase. The method is useful for producing host cells that express
CC recombinant proteins, which may be essential to the therapeutic treatment
CC of many diseases and conditions. The present sequence is retroviral
CC vector of the invention expressing hepatitis B surface antigen.
XX
SQ Sequence 5130 BP; 1187 A; 1335 C; 1332 G; 1276 T; 0 U; 0 Other;
Query Match 30.7%; Score 1588.6; DB 14; Length 5130;
Best Local Similarity 66.1%; Pred. No. 0;
Matches 3007; Conservative 0; Mismatches 754; Indels 786; Gaps 18;
QY 1 TTTGAAAGACCCACCCGTAGGTGGCAAGCTAGCTTAAGTAAACCCACTTTGCAAGGCAT 60
DB 1 TTTGAAAGACCCACCCGTAGGTGGCAAGCTAGCTTAAGTAAACCCACTTTGCAAGGCAT 60
QY 61 GGAATAATACATACCTAGAGATAGAAAGTTTCAGATCAAGTTCAGGAAACAAAGAACAGC 120
DB 61 GGAATAATACATACCTAGAGATAGAAAGTTTCAGATCAAGTTCAGGAAACAAAGAACAGC 120
QY 121 TGAATACCAACAGGATATCTGTGTAAGCGGTTCTGCCCCGGCTCAGGCCCAAGAAC 180
DB 121 TGAATACCAACAGGATATCTGTGTAAGCGGTTCTGCCCCGGCTCAGGCCCAAGAAC 180
QY 181 GATGAGACAGCTGAGTGATGGGCCAAACAGAGATATCTGTGTAAGCGGTTCTGCCCCGG 240
DB 181 GATGAGACAGCTGAGTGATGGGCCAAACAGAGATATCTGTGTAAGCGGTTCTGCCCCGG 240
QY 241 CTCGGGGCCCAAGAACAGATGTCCTCCAGATGCGGTCCAGCCCTCAGCAGTTTCTAGTGAA 300
DB 241 CTCGGGGCCCAAGAACAGATGTCCTCCAGATGCGGTCCAGCCCTCAGCAGTTTCTAGTGAA 300
QY 301 TCATCAGATGTTTCCAGGGTGCCTCCAGAGACCTGAAATGACCTTATTTGAAC 360
DB 301 TCATCAGATGTTTCCAGGGTGCCTCCAGAGACCTGAAATGACCTTATTTGAAC 360
QY 361 TAACCAATCAGTTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTT 420
DB 361 TAACCAATCAGTTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTT 420
QY 421 AGAGCCCAACACCCCTCCTCCTGCGGCGGCTTCTCCGATAGACTGCTGCGCGGGTAC 480
DB 421 AGAGCCCAACACCCCTCCTCCTGCGGCGGCTTCTCCGATAGACTGCTGCGCGGGTAC 480
QY 481 CCGTATTTCCTCAATAAAGCCTCTTGTCTTTGCTATCCGAATCGTGGTTCGCTGTTCTTGT 540
DB 481 CCGTATTTCCTCAATAAAGCCTCTTGTCTTTGCTATCCGAATCGTGGTTCGCTGTTCTTGT 540
QY 541 GGAGGGTCTCTCTGAGTGATGACTACCAACAGCGGGGTCTTTTCAATTTGGGGGCTCGT 600
DB 541 GGAGGGTCTCTCTGAGTGATGACTACCAACAGCGGGGTCTTTTCAATTTGGGGGCTCGT 600
```

QY 601 CCGGATTTGGAGACCCCTGCCAGGACACCCAGCCACCCAGGAGTAAGTGGCC 660
Db 601 CCGGATTTGGAGACCCCTGCCAGGACACCCAGCCACCCAGGAGTAAGTGGCC 660
QY 661 AGCAACTTATCTGTCTGTCGATGCTAGTGTCTATGTTGATGTTATGCGCTCG 720
Db 661 AGCAACTTATCTGTCTGTCGATGCTAGTGTCTATGTTGATGTTATGCGCTCG 720
QY 721 TCTGTACTAGTTAGTAACTAGTCTGTATCTGGCGGACCCGTGGTGAACCTGACGAGTT 780
Db 721 TCTGTACTAGTTAGTAACTAGTCTGTATCTGGCGGACCCGTGGTGAACCTGACGAGTT 780
QY 781 CTGAACACCCGCGCCCAACCTTGGAGAGCTCCAGGAGACTTTGGGCGCGTTTGTGG 840
Db 781 CTGAACACCCGCGCCCAACCTTGGAGAGCTCCAGGAGACTTTGGGCGCGTTTGTGG 840
QY 841 CCGGACCTGAGGAGGAGTGCATCTGGAATCCGACCCCGTCAGGATATGTTCTCGT 900
Db 841 CCGGACCTGAGGAGGAGTGCATCTGGAATCCGACCCCGTCAGGATATGTTCTCGT 900
QY 901 AGGAGACGAGAACCTTAAACAGTTCGCCCTCCGTCGTAATTTTTCGTTTGGAA 960
Db 901 AGGAGACGAGAACCTTAAACAGTTCGCCCTCCGTCGTAATTTTTCGTTTGGAA 960
QY 961 CCGAAGCCGCGGCTTGTCTGCTGCAGCAGAGCTTGGGCTGCAGGTCGACTTAGAGGA 1020
Db 961 CCGAAGCCGCGGCTTGTCTGCTGCAGCAGAGCTTGGGCTGCAGGTCGAGGACTGGGGA 1020
QY 1021 TCAATTCGGCAGGAGTAATCGGTGCTGCGCTCTTTAGGACATATGAGTATGACAGT 1080
Db 1021 CC-----CTGCACCGAACATGGAGAACACACATCAGGATTCCT 1059
QY 1081 GGGATGACTTTCTGATCAGCAAGAGGACACTGACAGCTGTACAGAGTCTGTGAAGTTCG 1140
Db 1060 AGGACCCCTGCTCGTGTACAGCGGGGTTTCTTGTGTGACAGAACTCTCAATACC 1119
QY 1141 ATGCTCGCTCAGTACAGCTTGTCTTCTCCCATCTTAAATAAGGCGCAACTCTTCAAG 1200
Db 1120 ACAGAGCTTAGAGCTCGTGTGAGTCTCTCAATTTCTAGGGGAGCACCAGCTGTCC 1179
QY 1201 AGAGGATGAAGTCTTATAAAGTGCATGATCACCCTTTATCTATGTTGTGTAGTTC 1260
Db 1180 TGGCCAAATTCGGAGTCCCAACCTCAATCACTCACCACCTCTTGTCTCCAAATTTG 1239
QY 1261 TCGTGCCCATCATTTGGCATGTGGCAGCTCAGCTCTGAAATGGGAAACGAGAAATGCA 1320
Db 1240 TCCTGGCTATC-----GCTGGATGTGTCTGGGCTTTATC 1276
QY 1321 CGGTTGGCTCAGTTAATCAGATATATCTCCAAGTCCGGAAGGCAAGAAATGGCAGTG 1380
Db 1277 ATATTCCTCTCATCTGCTGTATGCTCATCTTCTTGTGTTCTTCTGGACTACCA 1336
QY 1381 AAGATGAATAGATTTCCAGAGCTGTGATGGAACGATGAGCAACATGAGAAAGCAGAA 1440
Db 1337 GGTATGTTGCCGTTGTCTCTACTTCCAGGAA---CATCACTTACCAGCAGCGGACCA 1393
QY 1441 TCCAGTATCTTTCAGATAATGAAGCAATCTCTAGATGCTAAGAAATTTCCAAAATTTCA 1500
Db 1394 TGCAGAGCTGCAGATTTCTGCTCAAGGAACCTCTATGTTTCCCTCTT-----GTTGCT 1448
QY 1501 GCATTAACACTGATCAAAAGATTAATGATGTTCTTTTCCAGCTAAATCTTACTTTCCCT 1560
Db 1449 GTACAAAACCTTGGACGGAACCTGCACCTTGTATTTCCCATCCATCATCTTGGGCTTTCG 1508
QY 1561 CCATCCAGGAACATGAGATATCATAGGGGATATCTCCAAGTCATTAGTAGTCTGAACA 1620
Db 1509 CAAGATTCCTATGGAGTGGGCTCAGTCCGTTTCTCTCTCG----- 1549
QY 1621 CCACAGTACTTGTATTTGAGTTTCACTATTTGAACACTCAATGGCAGAGTCCAGAGATG 1680
Db 1550 -CTCAGTTTACTAGTCCATTTGTTTCACTGGTTC---GTAGGGCTTTCCCACTGTTTG 1605
QY 1681 CATTTTAAACAAAGAGGAGATGCGTAAATTTAGAGGAGCGTATATACAAATGCATCAGCAG 1740

Db 1606 GCTTTTCAGTTATATG-GATGATGTGGTATTGGGGCCCAAGTCTGTACAACTCTTGAGTC 1664
QY 1741 AAATTAAGTCTCTAGATGAAAAACAAGTATATTTTGGAAACAGGAAATAAAAGGGAATGA 1800
Db 1665 CCTTTTACCTCTATACCAATTTTCTTTTGTCTTTGGGTATACATTTAAACCCCTAATAA 1724
QY 1801 AACTGTTGAATAATATCACTAATGATCTGAGGCTGGAAGGATTTGGGAACATTTCTCAGACAT 1860
Db 1725 AAC-----CAACGCT 1734
QY 1861 TGAATAATATCACTTTACTCCAAAGTGCAGAAAGTGTCTGCTGACTGGGAAATGGACCA 1920
Db 1735 TGGGGCTACTCCCTTAACTTCATGGGATATGTAATTTGGATGTTGGGG----- 1781
QY 1921 ACATGCTGGGCTCAACATGACCATCGGGGCTGTGAACAGCAGAGGTAATTCACAGGCA 1980
Db 1782 ----- 1781
QY 1981 CCTACATCAGAGCGGTAAACAGCCACATCAAAATGAGATCAAAAGATCACCACTGCATGGGA 2040
Db 1782 -----TACTTTACCGCAAGAACATATTGTACTTAAAAATCAAGCAATGTTTTCG- 1829
QY 2041 CACAAAACACCATCAACAGAGGAGCCAGCCACCTTTGGGCTTCAACGTCATTTGGAAT 2100
Db 1830 -----AAACCTGCTGTAAATAGACCTATTGATTTGGAAGATGTGTAGAGACTTGTGGGT 1884
QY 2101 TTTCCAGAGTCCACCACTGTCTTTCAGCGGCCAGTCTTCATAGACAGGAATGGGAAGGAGG 2160
Db 1885 CTTTTGGGCTTGTGCTGCCCCCTTTTACAAATGTGGCTATCTGCTTAAATGCCCTTTATAT 1944
QY 2161 TCCTGAAGACATGTGGCTGCTGCGGTCAAGTGTAAATGACATTTGGTGTATGATGATGAAAG 2220
Db 1945 GCATGTATACATCT----- 1959
QY 2221 CTACAGAGTCCGATCAACATCTTCACTGCCCTGGCGCACAGAGAGTGTAGTGTG 2280
Db 1960 -----AAGCAGGCTTTTCACTTTCTCGCCAACTTACAAGGCTTTCTGTGTGA 2005
QY 2281 ACCAAGTCTCTCGGACTCCAGGTGAAAGAGGAGATAGAGGCCCTCTCGTGCACAAAATGG 2340
Db 2006 AACATATCTGAACCTTTACCCCGTGTCCCGGCA----- 2039
QY 2341 TATACAGGCTTTCCAGGTCTAATAGTGTCTTAAAGTGTCTTAAAGTGTCTCGGGGATCT 2400
Db 2040 ----- 2039
QY 2401 CTGTTTACCTGGAGTTGAGGATTTCCAGGACCAATGGGGAAGACCGGGAAGCCAGGAC 2460
Db 2040 ----- 2039
QY 2461 TTAATGGACAAAAGGCCAGAGGAGAAAAGGGAGTGGGAAGCATGCAAGAGACATCTA 2520
Db 2040 -----ACGTCAG 2047
QY 2521 ATACAGTCCGACTGGTGGTGGCGAGCGGCCCTCACGAAGGAGAGTGGAGATTTTTCACG 2580
Db 2048 GTCTCTGCCAAGTCTTTGTGACGCAACCCCACTGGATGGGCTTGGCTATCGGCCATA 2107
QY 2581 AAGGCCAGTGGGGTACGGTGTGTGACGACCGCTGGGAACTGCGGTGGAGGACTGCTGCTCT 2640
Db 2108 GCCGCATCGCGGACCTTTGTGGCTCTCTGCG----- 2141
QY 2641 GCAGGAGCTTGGGATACAAAGTGTCTTAAAGTGTGCATTAAGCAGCTTATTTTGGAAAAG 2700
Db 2142 -----ATCCATCTCGGAACTCCTAGCAGCTGT----- 2171
QY 2701 GTACGGTCCAAATATGCTGAATGAAGTATTTTCTCGGGAAGAGATCATCCATTTGAAG 2760
Db 2172 ----- 2171
QY 2761 AGTCAGAAATPAGACAGTGGGTTGTGAGAGCTGTTCGACGCAAGAGATCTCTGGGGTCT 2820

XX PD 22-JUL-1999.
XX PF 13-JAN-1999; 99WO-US000733.
XX PR 14-JAN-1998; 98US-0071409P.
XX PA (HUMA-) HUMAN GENE THERAPY RES INST.
XX PI Radosевич TJ, Link CJ;
XX DR WPI; 1999-468988/39.
XX PT Expression system containing therapeutic gene and an immunosuppressor
XX PT gene useful for treating an MHC-I autoimmune disease or killing tumor
XX PS cells.
XX CS Claim 25; Page 134-137; 154pp; English.
XX CC The present invention describes a nucleotide expression system for the
XX CC introduction of a therapeutic gene comprising: (i) a nucleotide sequence
XX CC encoding an immune suppression gene; (ii) a promoter; and (iii) a
XX CC transcription termination signal, where the system is able to inhibit,
XX CC evade or eliminate a recipient cell immune response to the therapeutic
XX CC gene when the gene is transformed into a recipient cell. AAX90481 to
XX CC AAX90484 represent specifically claimed plasmid retroviral vector
XX CC nucleotide sequences from the present invention. The expression system
XX CC and vectors containing it can be used for gene therapy, for treating an
XX CC MHC-I autoimmune disease or for killing tumor cells. The expression
XX CC system contains an immunosuppressive gene which prevents host rejection
XX CC of the vector
XX CS Sequence 6141 BP; 1404 A; 1698 C; 1603 G; 1436 T; 0 U; 0 Other;
SQ
Query Match 28.7%; Score 1486; DB 2; Length 6141;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3692 GATCTGATCAAGAGACAGGATGAGATCGTTTCGCATGATTGAACAAGATGGATTGCACG 3751
DB 2298 GATCTGATCAAGAGACAGGATGAGATCGTTTCGCATGATTGAACAAGATGGATTGCACG 2357
QY 3752 CAGGTTCTCCGGCCGCTTGGGTGGAGAGGCTATTTCGGCTATGACTTGGSCACAAACAGACAA 3811
DB 2358 CAGGTTCTCCGGCCGCTTGGGTGGAGAGGCTATTTCGGCTATGACTTGGSCACAAACAGACAA 2417
QY 3812 TCGGCTGCTCTGATGCGCCGCTGTTTCGGCTGTTCAGCGAGGGCGCCGGTCTCTTTTG 3871
DB 2418 TCGGCTGCTCTGATGCGCCGCTGTTTCGGCTGTTCAGCGAGGGCGCCGGTCTCTTTTG 2477
QY 3872 TCAGACCGACCTGTCCGGTGCCTTGATGAATGAACTGCAGGACGAGCGCGGCTATCGT 3931
DB 2478 TCAAGACCGACCTGTCCGGTGCCTTGATGAATGAACTGCAGGACGAGCGCGGCTATCGT 2537
QY 3932 GCCTGGCCACACGCGGCGTCTCTTGGCAGCTGTCTGCTGCTGCTCACTCAAGCGGAA 3991
DB 2538 GCCTGGCCACACGCGGCGTCTCTTGGCAGCTGTCTGCTGCTGCTCACTCAAGCGGAA 2597
QY 3992 GGGACTGGCTGCTATTGGCGCAAGTGCAGGCGGCGAGGATCTCTGTCACTCACCTTGCTC 4051
DB 2598 GGGACTGGCTGCTATTGGCGCAAGTGCAGGCGGCGAGGATCTCTGTCACTCACCTTGCTC 2657
QY 4052 CTGCCGAGAAAGTATCCATCATGGCTGATGCAATGCGCGCGGCTGATACGCTTGAATCCGG 4111
DB 2658 CTGCCGAGAAAGTATCCATCATGGCTGATGCAATGCGCGCGGCTGATACGCTTGAATCCGG 2717
QY 4112 CTACTCGCCATTTCGACCAACAGGAAACATTCGATCGAGCGAGCATCTCGATGG 4171
DB 2718 CTACTCGCCATTTCGACCAACAGGAAACATTCGATCGAGCGAGCATCTCGATGG 2777
QY 4172 AAGCCGGTCTTGTTCGATCAGGATGATCTGGACGAGAGCATCAGGGGCTCGCGCAGCGC 4231
DB 2778 AAGCCGGTCTTGTTCGATCAGGATGATCTGGACGAGAGCATCAGGGGCTCGCGCAGCGC 2837

QY 4232 AACTGTTCCGCGAGCTCAAGCGCGCATGCCGAGCGGAGGATCTCTGTCGTGACCCCATG 4291
DB 2838 AACTGTTCCGCGAGCTCAAGCGCGCATGCCGAGCGGAGGATCTCTGTCGTGACCCATG 2897
QY 4292 GCGATGCTGCTCTTCCCGAATATCATGTGGAAATAGCCCGCTTTTCTGGATTCATCGACT 4351
DB 2898 GCGATGCTGCTCTTCCCGAATATCATGTGGAAATAGCCCGCTTTTCTGGATTCATCGACT 2957
QY 4352 GTGGCCGGCTGGGTGGCGGACCGCTATCAGGACATAGCGTTCGCTACCGCTGATATTG 4411
DB 2958 GTGGCCGGCTGGGTGGCGGACCGCTATCAGGACATAGCGTTCGCTACCGCTGATATTG 3017
QY 4412 CTGAAGAGCTTGGCGGGAATGGCTGACCGCTTCTCTGCTGCTTTAGCGTATCCCGCTC 4471
DB 3018 CTGAAGAGCTTGGCGGGAATGGCTGACCGCTTCTCTGCTGCTTTAGCGTATCCCGCTC 3077
QY 4472 CCGATTCCGAGCGCATCGCTTCTATCGCTTCTTACGAGTCTTCTCTGAGCGGAGTCT 4531
DB 3078 CCGATTCCGAGCGCATCGCTTCTATCGCTTCTTACGAGTCTTCTCTGAGCGGAGTCT 3137
QY 4532 GGGTTTCGATAAAATAAAAGATTTTATTTAGTCTCCAGAAAAAGGGGGGAATGAAAGACC 4591
DB 3138 GGGTTTCGATAAAATAAAAGATTTTATTTAGTCTCCAGAAAAAGGGGGGAATGAAAGACC 3197
QY 4592 CCACCTGTAGTGTGGCAAGCTAGCTTAAGTAAAGCCATTTTGAAGGCAATGGAATAATA 4651
DB 3198 CCACCTGTAGTGTGGCAAGCTAGCTTAAGTAAAGCCATTTTGAAGGCAATGGAATAATA 3257
QY 4652 CATAACTGAGAAATAGAGAGTTTCAGATCAAGAGTTCAGGAAACAGATGGAACAGTGAATATG 4711
DB 3258 CATAACTGAGAAATAGAGAGTTTCAGATCAAGAGTTCAGGAAACAGATGGAACAGTGAATATG 3317
QY 4712 GGCCAAACAGGATATCTGTGGTAAGAGTTCCTGCCCCCGCTCAGGGCCAAAGAACAGATG 4771
DB 3318 GGCCAAACAGGATATCTGTGGTAAGAGTTCCTGCCCCCGCTCAGGGCCAAAGAACAGATG 3377
QY 4772 GACAGCTGAAATAGGCGCAACAGGATATCTGTGGTAAGAGTTCCTGCCCCCGCTCAG 4831
DB 3378 GACAGCTGAAATAGGCGCAACAGGATATCTGTGGTAAGAGTTCCTGCCCCCGCTCAG 3437
QY 4832 GGCCAAACAGGATGTTCCCGCAGATGCGGTCAGCCCTCAGCAGTTCCTAGAGAACCATC 4891
DB 3438 GGCCAAACAGGATGTTCCCGCAGATGCGGTCAGCCCTCAGCAGTTCCTAGAGAACCATC 3497
QY 4892 AGATGTTTCCAGGGTGGCCCAAGGACCTGAAATGACCTGTGCTTATTTGAACCTAACCA 4951
DB 3498 AGATGTTTCCAGGGTGGCCCAAGGACCTGAAATGACCTGTGCTTATTTGAACCTAACCA 3557
QY 4952 ATCAGTTTCGCTTCTCGCTTCTGTTTCGCGCGCTTCTGCTCCCGAGCTCAATAAAGAGCC 5011
DB 3558 ATCAGTTTCGCTTCTCGCTTCTGTTTCGCGCGCTTCTGCTCCCGAGCTCAATAAAGAGCC 3617
QY 5012 CACAACCCCTCAGTTCGCGGCGCAGTCTCTCCGATTCAGTTCGCGCGGTCACCCGTGT 5071
DB 3618 CACAACCCCTCAGTTCGCGGCGCAGTCTCTCCGATTCAGTTCGCGCGGTCACCCGTGT 3677
QY 5072 ATCCAAATAAACCTCTTTCGAGTTGCAATCCGACTTGTGGTCTCTGCTGTTTCTTGGAGGGT 5131
DB 3678 ATCCAAATAAACCTCTTTCGAGTTGCAATCCGACTTGTGGTCTCTGCTGTTTCTTGGAGGGT 3737
QY 5132 CTCTCTCTGAGTGAATGACTACCGCTCAGCGGGGCTCTTTTCATTTGG 5177
DB 3738 CTCTCTCTGAGTGAATGACTACCGCTCAGCGGGGCTCTTTTCATTTGG 3783

RESULT 14
AAX90482
ID AAX90482 standard; DNA; 6522 BP.
XX
XX AAX90482;
XX AC
XX AC
DT 29-SEP-1999 (first entry)

XX Plasmid retroviral vector pLUSN nucleotide sequence.
 XX DE
 XX Plasmid retroviral vector; expression system; immunogenic; gene therapy;
 KW immune response; immunosuppression; gene delivery; therapeutic;
 KW MHC-I autoimmune disease; tumour; ss.
 XX
 OS Synthetic.
 XX
 XX WO9936562-A1.
 XX
 XX 22-JUL-1999.
 PD
 XX 13-JAN-1999; 98WO-US000733.
 PF
 XX 14-JAN-1998; 98US-0071409P.
 PR
 XX (HUMA-) HUMAN GENE THERAPY RES INST.
 PA
 XX Radosevich TV, Link CJ;
 PI
 XX WPI; 1999-468988/39.
 DR
 XX Expression system containing therapeutic gene and an immunosuppressor
 PT gene useful for treating an MHC-I autoimmune disease or killing tumor
 PT cells.
 XX
 XX Claim 25; Page 137-140; 154pp; English.
 PS
 XX The present invention describes a nucleotide expression system for the
 CC introduction of a therapeutic gene comprising: (i) a nucleotide sequence
 CC encoding an immune suppression gene; (ii) a promoter; and (iii) a
 CC transcription termination signal, where the system is able to inhibit,
 CC evade or eliminate a recipient cell immune response to the therapeutic
 CC gene when the gene is transformed into a recipient cell. AA90481 to
 CC AA90484 represent specifically claimed plasmid retroviral vector
 CC nucleotide sequences from the present invention. The expression system
 CC and vectors containing it can be used for gene therapy, for treating an
 CC MHC-I autoimmune disease or for killing tumour cells. The expression
 CC system contains an immunosuppressive gene which prevents host rejection
 CC of the vector
 CC
 XX
 XX SQ Sequence 6522 BP; 1469 A; 1792 C; 1704 G; 1557 T; 0 U; 0 Other;
 Query Match 28.7%; Score 1486; DB 2; Length 6522;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 3692 GATCTGATCAAGAGACAGAGATGAGATCGTTTCGATGATTGAAACAGATGGAATGCAAG 3751
 2679 GATCTGATCAAGAGACAGAGATGAGATCGTTTCGATGATTGAAACAGATGGAATGCAAG 2738
 3752 CAGGTTCTCCGGCGCTTGGGTGAGAGGCTATTTCGGCTATGACTGGGCACACAGACAA 3811
 2739 CAGGTTCTCCGGCGCTTGGGTGAGAGGCTATTTCGGCTATGACTGGGCACACAGACAA 2798
 3812 TCGGCTGCTCTGATGCGCGCGTGTTCGGCTGTGAGCGCGCGCGCGCGTCTTTTGG 3871
 2799 TCGGCTGCTCTGATGCGCGCGTGTTCGGCTGTGAGCGCGCGCGCGCGTCTTTTGG 2858
 3872 TCAAGACCGAAGCTGTCGGTCCCTGGAATGAATGACAGAGACGAGCGCGGTATCGT 3931
 2859 TCAAGACCGAAGCTGTCGGTCCCTGGAATGAATGACAGAGACGAGCGCGGTATCGT 2918
 3932 GGCTGGCCACAGCGGGCGTCTTTCGGCGAGCTGTGCTCGAGCTTGTCACTGAACGGGAA 3991
 2919 GGCTGGCCACAGCGGGCGTCTTTCGGCGAGCTGTGCTCGAGCTTGTCACTGAACGGGAA 2978
 3992 GGGACTGCTGCTATTGGGCGAAGTGGCGGCGAGGATCTCTGTGCTCACTCACTTGTCTC 4051
 2979 GGGACTGCTGCTATTGGGCGAAGTGGCGGCGAGGATCTCTGTGCTCACTCACTTGTCTC 3038
 4052 CTGCCGAGAAAGTATCCATCATGCTGATGCAATGCGCGCGCTGCATACGCTTGTATCGG 4111

Db 3039 CTGCCGAGAAAGTATCCATCATGCTGATCAATGCGCGGCTGCATACGCTTATCCGG 3098
 Qy 4112 CTACTCTCCCATTCGACCCACCAAGCGAAACATCCCATCGAGCGAGCAGTACTCGGATGG 4171
 Db 3099 CTACTCTCCCATTCGACCCACCAAGCGAAACATCCCATCGAGCGAGCAGTACTCGGATGG 3158
 Qy 4172 AAGCCGCTCTTGTCTGATCAGATGATCTGAGCAAGAGCATACAGGGGCTCGCGCAGCG 4231
 Db 3159 AAGCCGCTCTTGTCTGATCAGATGATCTGAGCAAGAGCATACAGGGGCTCGCGCAGCG 3218
 Qy 4232 AACTGTTGGCCAGGCTCAAGCGCGCATGCGCGAGGATCTCGTCTGAGCCCATG 4291
 Db 3219 AACTGTTGGCCAGGCTCAAGCGCGCATGCGCGAGGATCTCGTCTGAGCCCATG 3278
 Qy 4292 GCGATGCTCTTGCCTCGAATATCATGCTGGAATGCGCGCTTTCTTGATTCATCGACT 4351
 Db 3279 GCGATGCTCTTGCCTCGAATATCATGCTGGAATGCGCGCTTTCTTGATTCATCGACT 3338
 Qy 4352 GTGGCCGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCGCTGATATG 4411
 Db 3339 GTGGCCGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCGCTGATATG 3398
 Qy 4412 CTGAGAGACTTGGCGGCAATGGGCTGACCGCTTCTCGTCTTACGGTATCGCGCTC 4471
 Db 3399 CTGAGAGACTTGGCGGCAATGGGCTGACCGCTTCTCGTCTTACGGTATCGCGCTC 3458
 Qy 4472 CCGATTCCGAGCGCATCGCTTCTATCGCTTCTTGACAGTCTTCTTGACGCGGACTCT 4531
 Db 3459 CCGATTCCGAGCGCATCGCTTCTATCGCTTCTTGACAGTCTTCTTGACGCGGACTCT 3518
 Qy 4532 GGGGTTTCGATAAAATAAAAGATTTTATTAGTCTCCAGAAAAAGGGGGAATGAAAGACC 4591
 Db 3519 GGGGTTTCGATAAAATAAAAGATTTTATTAGTCTCCAGAAAAAGGGGGAATGAAAGACC 3578
 Qy 4592 CCACCTTAGCTTGGCAAGCTAGCTTAAGTAAGCCATTTTGCAGGCAATGGAATAATA 4651
 Db 3579 CCACCTTAGCTTGGCAAGCTAGCTTAAGTAAGCCATTTTGCAGGCAATGGAATAATA 3638
 Qy 4652 CATAACTGAGAAATAGAGAGTTTACAGATCAAGGTCAGGAACAGATGGAACAGCTCAATATG 4711
 Db 3639 CATAACTGAGAAATAGAGAGTTTACAGATCAAGGTCAGGAACAGATGGAACAGCTCAATATG 3698
 Qy 4712 GGCCAAACAGGATATCTGTGTAGAGTTCCTGCGCGGCTCAGGGCCAGAAACAGATG 4771
 Db 3699 GGCCAAACAGGATATCTGTGTAGAGTTCCTGCGCGGCTCAGGGCCAGAAACAGATG 3758
 Qy 4772 GAACAGCTGAATATGGGCCAAAACAGGATATCTGTGTAGAGTTCCTGCGCGGCTCAG 4831
 Db 3759 GAACAGCTGAATATGGGCCAAAACAGGATATCTGTGTAGAGTTCCTGCGCGGCTCAG 3818
 Qy 4832 GGCCAAACAGAGATGTCCTCCAGATGCGGTCCAGCCCTCAGCAGTTCCTAGAGAACCATC 4891
 Db 3819 GGCCAAACAGAGATGTCCTCCAGATGCGGTCCAGCCCTCAGCAGTTCCTAGAGAACCATC 3878
 Qy 4892 AGATGTTTCCAGGGTGCCCAAGGACCTGAAATGACCTGTGCTTATTTGAACTAACCA 4951
 Db 3879 AGATGTTTCCAGGGTGCCCAAGGACCTGAAATGACCTGTGCTTATTTGAACTAACCA 3938
 Qy 4952 ATCAGTTTCGCTTCTCGCTTCTGTTCGCGCGCTTCTGCTCCCGAGCTCAATAAAGAGCC 5011
 Db 3939 ATCAGTTTCGCTTCTCGCTTCTGTTCGCGCGCTTCTGCTCCCGAGCTCAATAAAGAGCC 3998
 Qy 5012 CACAACCCCTCAGTCGGGGCGCAGTCTCCGATGAGTTCAGTCCCGCGGCTACCCGTGT 5071
 Db 3999 CACAACCCCTCAGTCGGGGCGCAGTCTCCGATGAGTTCAGTCCCGCGGCTACCCGTGT 4058
 Qy 5072 ATCCAAATAAACCCCTCTTTCGAGTTGCAATCCGACTTGTGCTCTCGCTTCTTGGAGGGT 5131
 Db 4059 ATCCAAATAAACCCCTCTTTCGAGTTGCAATCCGACTTGTGCTCTCGCTTCTTGGAGGGT 4118
 Qy 5132 CTCTCTGAGTGAATGATACCCGCTCAGCGGGGTCTTTTCATTTGG 5177

| | | | |
|-----------|---|---|------|
| Db | 4119 | CTCCTCTGAGTGATTGACTACCGCTGACGGGGGGTCTTTCATTTGG | 4164 |
| RESULT 15 | | | |
| ADBE82653 | | | |
| ID | ADBE82653 | standard; DNA; 6620 BP. | |
| AC | ADBE82653; | | |
| XX | | | |
| DT | 29-JAN-2004 | (first entry) | |
| XX | | | |
| XX | | Plasmid vector pLESN nucleotide sequence SEQ ID NO:3. | |
| XX | | | |
| KW | | recombinant DNA construct; | |
| KW | | humanised red shifted green fluorescent protein; HRGFP; fluorescence; | |
| KW | | gene transfer; gene expression; gene therapy; bone marrow processing; | |
| KW | | plasmid vector; gene; ds. | |
| XX | | | |
| OS | | Synthetic. | |
| XX | | | |
| PN | US2002015979-A1. | | |
| XX | | | |
| PD | 07-FEB-2002. | | |
| XX | | | |
| PF | 21-JAN-1997; | 97US-00786531. | |
| XX | | | |
| PR | 22-JAN-1996; | 96US-0010371P. | |
| XX | | | |
| XX | | (LINK/) LINK C J. | |
| PA | (LEVI/) LEVI J P. | | |
| PA | (WANG/) WANG S. | | |
| PA | (SERE/) SEREGINA T. | | |
| XX | | | |
| PI | Link CJ, Levy JP, Wang S, Seregina T; | | |
| XX | | | |
| DR | WPI; 2002-239224/29. | | |
| XX | | | |
| PT | Recombinant DNA construct for identifying transformed cells to allow for | | |
| PT | direct observation of transferred genes into living cells, comprising | | |
| PT | nucleotide sequence encoding red shifted green fluorescent protein. | | |
| XX | | | |
| PS | Claim 17; SEQ ID NO 3; 95pp; English. | | |
| XX | | | |
| CC | The present invention describes a recombinant DNA construct (I), | | |
| CC | comprising a nucleotide sequence encoding upon expression a humanised red | | |
| CC | shifted green fluorescent protein (HRGFP), and a promoter sequence and a | | |
| CC | termination sequence functionally coupled to the coding sequence. Also | | |
| CC | described: (1) a plasmid vector DNA sequence (II) enabling replication of | | |
| CC | the vector or in a host cell and (I); (2) infectious virus (III) | | |
| CC | comprising retroviral RNA transcribed from (I) in a host cell capable of | | |
| CC | viral packaging; and (3) a mammalian cell (IV) or a cell derived from it, | | |
| CC | comprising at least one copy of (I). (I) is useful for identifying | | |
| CC | transformed cells to allow for direct observation of transferred genes | | |
| CC | into living cells, by introducing (I) into the cell and measuring | | |
| CC | fluorescence of HRGFP, expressed by the transformed cells. The | | |
| CC | fluorescence is measured by the absorbance at 490-760 nm, and the | | |
| CC | excitation is from 420-470 nm. (I) is also useful for selecting | | |
| CC | lymphocytes which have been transfected with a vector, by transforming | | |
| CC | cells with (I) and sorting the cells with a fluorescence activated cell | | |
| CC | sorter (FACS). The expression of the HRGFP gene in living eukaryotic | | |
| CC | cells will advance the study of gene transfer, gene expression and gene | | |
| CC | product function in vitro or in vivo, particularly for human gene therapy | | |
| CC | applications. (I) is useful in bone marrow processing and other | | |
| CC | applications requiring FACS analyses of living cells. The present | | |
| CC | sequence represents a specifically claimed plasmid vector nucleotide | | |
| CC | sequence, which is used in the exemplification of the present invention. | | |
| CC | N.B. The SEQ ID NO:s in the Sequence Listing do not correspond with SEQ | | |
| CC | ID NO:s mentioned in the specification: where SEQ ID NO:1 to 6 of the | | |
| CC | Sequence Listing represent the sequences in figures 22, 17, 18, 20, 19 | | |
| CC | and 21 respectively; SEQ ID NO:7 to 10 of the Sequence Listing represents | | |
| CC | SEQ ID NO:2 to 5 on page 7 of example 1; and SEQ ID NO:1 of the | | |
| CC | specification given in figure 7 is not given in the Sequence Listing. | | |
| XX | | | |

QY 4712 GGCCAAACAGGATATCTGTGGTAAGCAGTTCTGCCCCGGCTCAGGGCCAAAGAACAGATG 4771
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3797 GGCCAAACAGGATATCTGTGGTAAGCAGTTCTGCCCCGGCTCAGGGCCAAAGAACAGATG 3856
QY 4772 GAAACAGCTGAATATGGGCCAAACAGGATATCTGTGGTAAGCAGTTCTGCCCCGGCTCAG 4831
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3857 GAAACAGCTGAATATGGGCCAAACAGGATATCTGTGGTAAGCAGTTCTGCCCCGGCTCAG 3916
QY 4832 GGCCAAAGAACAGATGGTCCCCCAGATGCGGTCCAGCCCTCAGCAGTTTCTAGAGAACCATC 4891
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3917 GGCCAAAGAACAGATGGTCCCCCAGATGCGGTCCAGCCCTCAGCAGTTTCTAGAGAACCATC 3976
QY 4892 AGATGTTTCCAGGGTGCCCCAAAGAACCTGAATGACCTGTGCTTATTTGAACTAAACA 4951
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3977 AGATGTTTCCAGGGTGCCCCAAAGAACCTGAATGACCTGTGCTTATTTGAACTAAACA 4036
QY 4952 ATCAGTTCGCTTCTCGCTTCTGTTGCGGCGCTTCTGCTCCCGAGCTCAATAAAGAGCC 5011
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
4037 ATCAGTTCGCTTCTCGCTTCTGTTGCGGCGCTTCTGCTCCCGAGCTCAATAAAGAGCC 4096
QY 5012 CACAAACCCCTCACTCGGGGCGCCAGTCTCCGATTGACTGAGTCGCCGGGTACCCGTGT 5071
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
4097 CACAAACCCCTCACTCGGGGCGCCAGTCTCCGATTGACTGAGTCGCCGGGTACCCGTGT 4156
QY 5072 ATCCAAATAAACCCCTTTGCACTTGCATCCGACTTGTGCTCTCGCTGTTCTTGGAGGGT 5131
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
4157 ATCCAAATAAACCCCTTTGCACTTGCATCCGACTTGTGCTCTCGCTGTTCTTGGAGGGT 4216
QY 5132 CTCCTCTGAGTGATTGACTACCCGTACGCGGGGTCTTTCAATTGG 5177
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
4217 CTCCTCTGAGTGATTGACTACCCGTACGCGGGGTCTTTCAATTGG 4262

Search completed: February 27, 2006, 09:24:01
Job time : 2890 secs

THIS PAGE BLANK (USPTO)

QY 181 GATGAGACAGCTGAGTGTGGCCCAACAGAGATATCTGTGTAAAGCAGTCTCTGCCCCG 240
DB |||||
QY 181 GATGAGACAGCTGAGTGTGGCCCAACAGAGATATCTGTGTAAAGCAGTCTCTGCCCCG 240
DB |||||
QY 241 CTGGGGCCAGAGAACAGATGTGCCAGATGGGTCCAGCCCTCAGCAGTTTCTAGTGAA 300
DB |||||
QY 241 CTGGGGCCAGAGAACAGATGTGCCAGATGGGTCCAGCCCTCAGCAGTTTCTAGTGAA 300
DB |||||
QY 301 TCATCAGATGTTTCCAGGGTCCCAAGGACCTGAAATGACCTGTACCTTATTGNAAC 360
DB |||||
QY 301 TCATCAGATGTTTCCAGGGTCCCAAGGACCTGAAATGACCTGTACCTTATTGNAAC 360
DB |||||
QY 361 TAAACCAATCAGTCTCGCTTCTGTCTGTCTGCGCGCTTCCGCTCTCCGAGCTCAATAAA 420
DB |||||
QY 361 TAAACCAATCAGTCTCGCTTCTGTCTGTCTGCGCGCTTCCGCTCTCCGAGCTCAATAAA 420
DB |||||
QY 421 AGAGCCCAACCCCTCCTGCGGCGCCAGCTTCCGATAGACTGCGTCCCGGGTAC 480
DB |||||
QY 421 AGAGCCCAACCCCTCCTGCGGCGCCAGCTTCCGATAGACTGCGTCCCGGGTAC 480
DB |||||
QY 481 CCGTATTTCCCAATAAGCCCTCTGTGTTTGCATCCGAATCGTGTCTCGCTGTCTCTTG 540
DB |||||
QY 481 CCGTATTTCCCAATAAGCCCTCTGTGTTTGCATCCGAATCGTGTCTCGCTGTCTCTTG 540
DB |||||
QY 541 GGAGGGTCTCTCTGAGTGATTGACTACCCACGACGGGGTCTTTTCAATTTGGGGCTCGT 600
DB |||||
QY 541 GGAGGGTCTCTCTGAGTGATTGACTACCCACGACGGGGTCTTTTCAATTTGGGGCTCGT 600
DB |||||
QY 601 CCGGATTTGAGACCCCTGCGGAGGACCAACGACCCACCGGGAGGTAAAGCTGCGC 660
DB |||||
QY 601 CCGGATTTGAGACCCCTGCGGAGGACCAACGACCCACCGGGAGGTAAAGCTGCGC 660
DB |||||
QY 661 AGCAACTTATCTGTGTCTGTCGATGTCTAGTGTCTATGTTGATTTAGCGCGCTGCG 720
DB |||||
QY 661 AGCAACTTATCTGTGTCTGTCGATGTCTAGTGTCTATGTTGATTTAGCGCGCTGCG 720
DB |||||
QY 721 TCTGTACTAGTTAGCTAACTAGCTCTGTATCTGGCGGACCGTGTGGGAACTGACGAGTT 780
DB |||||
QY 721 TCTGTACTAGTTAGCTAACTAGCTCTGTATCTGGCGGACCGTGTGGGAACTGACGAGTT 780
DB |||||
QY 781 CTGAACACCCGCGCAACCCCTGGGAGACGTCCCAAGGACTTTGGGGCCCGTTTTTGTG 840
DB |||||
QY 781 CTGAACACCCGCGCAACCCCTGGGAGACGTCCCAAGGACTTTGGGGCCCGTTTTTGTG 840
DB |||||
QY 841 CCGGACCTGAGGAGGAGTGAATGTGAATCCGACCCCGTCAAGATATGTTGTTCTGCT 900
DB |||||
QY 841 CCGGACCTGAGGAGGAGTGAATGTGAATCCGACCCCGTCAAGATATGTTGTTCTGCT 900
DB |||||
QY 901 AGGAGACGAGAACCTTAAACAGTTCCTCGCCTCCGTCTGAAATTTTTCGTTTGGAA 960
DB |||||
QY 901 AGGAGACGAGAACCTTAAACAGTTCCTCGCCTCCGTCTGAAATTTTTCGTTTGGAA 960
DB |||||
QY 961 CCGAAGCCGCGCTTGTCTGCTGACGAGCTTGGGCTGCAAGTCTGAGGAACTCTAGAGGA 1020
DB |||||
QY 961 CCGAAGCCGCGCTTGTCTGCTGACGAGCTTGGGCTGCAAGTCTGAGGAACTCTAGAGGA 1020
DB |||||
QY 1021 TCNATTCGACAGTAAATCGGTCTCGCTCTTTAGGACATATGAAGTATGACAGT 1080
DB |||||
QY 1021 TCNATTCGACAGTAAATCGGTCTCGCTCTTTAGGACATATGAAGTATGACAGT 1080
DB |||||
QY 1081 GGGATGACTTTTCTGATCAGCAAGAGGACACTGACAGCTGTACAGAGTCTGTGAAGTTTCG 1140
DB |||||
QY 1081 GGGATGACTTTTCTGATCAGCAAGAGGACACTGACAGCTGTACAGAGTCTGTGAAGTTTCG 1140
DB |||||
QY 1141 ATGCTCGCTCAGTGACAGCTTTTGTCTCTCCCATCTTAAATAATGGCCCAACTCTTCAAG 1200
DB |||||
QY 1141 ATGCTCGCTCAGTGACAGCTTTTGTCTCTCCCATCTTAAATAATGGCCCAACTCTTCAAG 1200
DB |||||
QY 1201 AGAGATGAAGTCTTATATAAACTGACATGATCAACCTTTATCTCATTTGTGTTGTAGTTC 1260
DB |||||
QY 1201 AGAGATGAAGTCTTATATAAACTGACATGATCAACCTTTATCTCATTTGTGTTGTAGTTC 1260
DB |||||
QY 1261 TCGTGCCCATCATTTGGCATAGTGGCAGCTCAGCTCCTGAAATGGGAAACGAAGATTGCA 1320

DB ||||| TCGTGCCCATCATTTGGCATAGTGGCAGCTCAGCTCCTGNAATGGGAAACGAAGATTGCA 1320
QY |||||
QY 1321 CGGTTGGCTCAGTTAAATGCGAGATATATCTCAAGTCCGGAAGGCAAGAAATGGCAGTG 1380
DB |||||
QY 1321 CGGTTGGCTCAGTTAAATGCGAGATATATCTCAAGTCCGGAAGGCAAGAAATGGCAGTG 1380
DB |||||
QY 1381 AAGATGAAATGAGATTTTCGAGAGAGCTGTGATGGAAACGATGAGCAACATGGAAGCAGAA 1440
DB |||||
QY 1381 AAGATGAAATGAGATTTTCGAGAGAGCTGTGATGGAAACGATGAGCAACATGGAAGCAGAA 1440
DB |||||
QY 1441 TCCAGTATCTTTTTCAGATAATGAAGCCAAATCTCTCTAGATGCTTAAAGATTTTCCAAATTTCA 1500
DB |||||
QY 1441 TCCAGTATCTTTTTCAGATAATGAAGCCAAATCTCTCTAGATGCTTAAAGATTTTCCAAATTTCA 1500
DB |||||
QY 1501 GCATPAAACATGTATCAAGATTTTAAATGATGTTCTTTTCCAGCTAAATCTCTTACTTCTCT 1560
DB |||||
QY 1501 GCATPAAACATGTATCAAGATTTTAAATGATGTTCTTTTCCAGCTAAATCTCTTACTTCTCT 1560
DB |||||
QY 1561 CCATCCAGGAAACATGAGATATCATAGGGATATCTCCAAAGTCAATAGTAGGTCTGAACA 1620
DB |||||
QY 1561 CCATCCAGGAAACATGAGATATCATAGGGATATCTCCAAAGTCAATAGTAGGTCTGAACA 1620
DB |||||
QY 1621 CCAGTACTTGTATTTGCGAGTTTCAAGTATGAAAACACTGAAATGGCAGAGTCCAAGAGAAATG 1680
DB |||||
QY 1621 CCAGTACTTGTATTTGCGAGTTTCAAGTATGAAAACACTGAAATGGCAGAGTCCAAGAGAAATG 1680
DB |||||
QY 1681 CATTTAAACAAAGAGGAGATGCGTAAATAGAGGAGCGTATATACAAATGCATCAGCAG 1740
DB |||||
QY 1681 CATTTAAACAAAGAGGAGATGCGTAAATAGAGGAGCGTATATACAAATGCATCAGCAG 1740
DB |||||
QY 1741 AAATTAAGTCTCTAGATGAAAACAACTATATTTGGAAACAGGAAATAAAAGGGGAAATGA 1800
DB |||||
QY 1741 AAATTAAGTCTCTAGATGAAAACAACTATATTTGGAAACAGGAAATAAAAGGGGAAATGA 1800
DB |||||
QY 1801 AACTGTTTGAATTAATCATTAATGATCTGAGGCTGAAAGGATTTGGGAAACATCTCTCAGACAT 1860
DB |||||
QY 1801 AACTGTTTGAATTAATCATTAATGATCTGAGGCTGAAAGGATTTGGGAAACATCTCTCAGACAT 1860
DB |||||
QY 1861 TGAATAATATCATTTTACTCCAAAGTCCCAAGAAAGTCTCGCTGACTGGGAAATGGACCA 1920
DB |||||
QY 1861 TGAATAATATCATTTTACTCCAAAGTCCCAAGAAAGTCTCGCTGACTGGGAAATGGACCA 1920
DB |||||
QY 1921 ACGATCTGGGCTCCAACTGACATGACCATCGGGCTGTGAACAGCAGAGGTGAATTCACAGCA 1980
DB |||||
QY 1921 ACGATCTGGGCTCCAACTGACATGACCATCGGGCTGTGAACAGCAGAGGTGAATTCACAGCA 1980
DB |||||
QY 1981 CCTACATCAGAGCGGTAAACAGCCACATCAAAATGAGATCAAAAGAGTCAACACTGATGGGA 2040
DB |||||
QY 1981 CCTACATCAGAGCGGTAAACAGCCACATCAAAATGAGATCAAAAGAGTCAACACTGATGGGA 2040
DB |||||
QY 2041 CACAAAACACATCAACAGAGGACCCAGCCCACTTTGGCTTCCCGCTCAATTTGGAGT 2100
DB |||||
QY 2041 CACAAAACACATCAACAGAGGACCCAGCCCACTTTGGCTTCCCGCTCAATTTGGAGT 2100
DB |||||
QY 2101 TTTTCAGAGTCCACACTGCTTCCAGGCGCAGTCTTATAGACAGGAAATGGGAGGAGG 2160
DB |||||
QY 2101 TTTTCAGAGTCCACACTGCTTCCAGGCGCAGTCTTATAGACAGGAAATGGGAGGAGG 2160
DB |||||
QY 2161 TCCTGAAAGACATGTTGCTGCTCGGCTCAAGTGTAAATGACATTTGGTGAATGGAAAG 2220
DB |||||
QY 2161 TCCTGAAAGACATGTTGCTGCTCGGCTCAAGTGTAAATGACATTTGGTGAATGGAAAG 2220
DB |||||
QY 2221 CTACAGGGTCCGATCAACATCTTCTACTCGCTGCGCACACAGAGAGTGAAGTGAAGT 2280
DB |||||
QY 2221 CTACAGGGTCCGATCAACATCTTCTACTCGCTGCGCACACAGAGAGTGAAGTGAAGT 2280
DB |||||
QY 2281 ACCAAGTCTCTGAGTCCAGGTGAAAAGAGATAGAGGCCCTCTCTGACAAAATGG 2340
DB |||||
QY 2281 ACCAAGTCTCTGAGTCCAGGTGAAAAGAGATAGAGGCCCTCTCTGACAAAATGG 2340
DB |||||
QY 2341 TATACAGGCTTTTCCAGTCTTAATAGGTACTCAGGCTTTAAAGGTGATCGGGGGGATCT 2400
DB |||||

1141 ATGCTGCTCAGTGACAGCTTTGCTTCTCTCCCAATCTCTAAATAATGGGCCCAACTCTTCAAG 1200
1120 ACAGAGTCTAGACTCGTGTGGAGCTTCTCTCAATTTCTAGGGGAGCACCACGCTGTCC 1179
1201 AGAGGATGAAGTCTTATAAACTGCACGTGATCACCTTTTATCTCATTTGTTGTAGTTC 1260
1180 TGGCCAAAATTCGCGAGTCCCAACTCCCACTCACTACCAACTCTTGTCTCTCAATTTG 1239
1261 TCGTGCCCATCAITTTGGCATAGTGGCAGCTCAGCTCTGAAATGGGAAACGGAATTCGA 1320
1240 TCCTGGCTATC-----GCTGGATGTCTCTGGCGGTTTATC 1276
1321 CGGTTGGCTCAGTTAATGCGAGATATCTCAGAGTCCGGAAGGCAAGGAATGGCAGTG 1380
1277 ATATTTCTCTTCTCCTCTGCTATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1336
1381 AAGATGAATCAGATTTTCGAGAGCTGTGAGGACGATGAGCAATGAGCAATGAGGAGCA 1440
1337 GGTATGTTGGCCGTTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1393
1441 TCCAGTATCTTTCCAGATAATGAAGCAATCTCTAGATGCTAAGAAATTTCCAAAATTTCA 1500
1394 TGGAGACCTGCAAGATTCCTGCTCAAGAACTCTATGTTTCCCTCTT-----GTTGCT 1448
1501 GCATAACAACTGATCAAAAGATTTAATGATGTTCTTTTCCAGCTAAATTTCTTACTTCT 1560
1449 GTACAAAACCTTCGAGCGAAACTGCACCTGTATTTCCCATCCATCCCTCTGGGCTTGG 1508
1561 CCATCCAGGAACATGAGATATATAGGGATATCTCCAGTCAATAGTAGGTCTGAACA 1620
1509 CAGATTTCTATGGAGTGGCCCTCAGTCCGTTTCTCTCTG----- 1549
1621 CCACAGTACTTGATTTGAGTTTCAAGTTCAGTATGAAACACTGAATGGGAGAGTCCAAAGAA 1680
1550 -CTCAGTTTACAGTGGCATTTGTTTCAAGTGTTC---GTAGGGCTTTTCCCACTGTTG 1605
1681 CATTAAACAAAGAGGAGATGCGTAAATPAGAGGAGCGTATATACAAATGATCATCAGCAG 1740
1606 GCTTTCAAGTATATG-GATGATGTTGATTTGGGGCCCAAGTCTGTACAACTCTTGAGTC 1664
1741 AAATTAAGTCTAGATGAAGAAACAAAGTATATTTGGAACAGAAATAAAGGGGAATGA 1800
1665 CTTTTTACCTCTATACCAATTTTCTTTGCTTTGGGTATACATTTAAACCCCTAATA 1724
1801 AACTGTTGAATAATCACTAATGATCTGAGGCTGAGGATTTGGAAACATTTCTCAGACAT 1860
1725 AAC-----CAAGCT 1734
1861 TGAATAATATCACTTTTACTCCAAGTGCAGAAAGTGTCTGCTGACTGGGAAATGGACCA 1920
1735 TGGGGCTACTCCCTTAACCTTCAAGGATATGTAATGGATGTTGGG----- 1781
1921 ACGATCTGGGCTCCAAATGACATCGGGCTGTGAACAGCAGAGGTTGAATTCACAGGCA 1980
1782 ----- 1781
1981 CCTATATCACAGCGTAAACGCCACATCAATGAGATCAAGAGTCAACCTGATGGGA 2040
1782 -----TACTTTTACCCGAGAACATATGTAATAAATCAAGCAATGTTTTCG- 1829
2041 CACAAAACCATCAACAAAGAGGCCACCGCTTTGGCTTCCCGCTCAATTTGGAAGT 2100
1830 -----AAAACTGCTGTAAATAGACCTATTGATTTGGAAGTATGTCAGAGACTTGTGGT 1884
2101 TTTGAGTCCACCATGCTCTTCAAGGCCAGTGTCTCATAGACAGGAATGGGAAGG 2160
1885 CTTTGGGCTTTGTGCGCCCTTTTACAAATGTGGCTATCTGCTTAATGCCCTTTATAT 1944
2161 TCCTGAGACCATGTGGCTGTGGCTCAAGTGTAAATGACATTTGGTGATGACTGGAAG 2220
1945 GCATGTATACAACTCT----- 1959

2221 CTACCAGGTCGGGATCAACATCTTCACTCCCTGCGCACACACAGAGTGAAGTG 2280
1960 -----AAGCAGGCTTTCACTTTCTCGCAACTTACAGGCTTTCTGTGTA 2005
2281 ACCAAGTCTCTCTGAGCTCCAGGTGAAAAGAGAGATAGAGGCCCTCTCTGGACAAAATGG 2340
2006 AACATATCTGAACCTTTACCCGTTGCCGGCA----- 2039
2341 TATACCAGGCTTTCCAGGCTCTAATAGGTACTCCAGGCTTTAAAGGTGATCGGGGGATCT 2400
2040 ----- 2039
2401 CTGGTTTACCTGGAGTTCGAGGATCCAGAGCAATGGGGAGACCGGAGCCAGGAC 2460
2040 ----- 2039
2461 TTAATGACAAAAGGCCAGAGGAGGAGAAAAGGGAGTGGGAAGCATGCAAGACAAATCTA 2520
2040 -----ACGTCAG 2047
2521 ATACAGTCCGACTGGTGGTGCGCAGCGCCCTCAAGAGGAGAGTGGAGATTTTTCAGG 2580
2048 GTCTCTGCAAGTGTGCTGACGCAACCCCACTGCTGGGCTTGGCTATCGGCCATA 2107
2581 AAGCCAGTGGGGTACGCTGTGTGACGACCGCTGGGAACTGCGTGGAGGACTGGTCTCT 2640
2108 GCGCATGCGCGGACCTTTGTGGCTCTCTGCGC----- 2141
2641 GCAGGACTTGGGATCAAAAGGTGTTCAAGGTGTCATTAAGCAGCTTATTTTGGAAAAG 2700
2142 -----ATCCATATCTGCGAACTCTTAGCAGCTTGT----- 2171
2701 GTACGGGTCCAAATATGCTGAATGAAGTATTTTGTTCGGGAAGAGATCATCTCAATTTGAAG 2760
2172 ----- 2171
2761 AGTGCAGAAATTAGACAGTGGGTGTGAGAGCTGTCGACGACGAAGATGCTCGGGGTC 2820
2172 -----TTTGCTCGCAGGCGGTCTGGAGCGAA 2197
2821 ACTTTGACCTACATATGATCATATTTTTCATTCACATTTTAAACTGTTATAAGTG 2880
2198 ACTTATCGGC----- 2207
2881 ATTTTTTCTTGTCTTCACTAAATCAGCTTAATTAATTTAAGAAAATAAGAAATTTT 2940
2208 ----- 2207
2941 ATCCACAGAAAAGGAATATTTAAATAATCATCTGGATAAACATATAAATAGCTTCATATTT 3000
2208 -----ACCGACAATCTCT 2219
3001 GCTTCAATACCAGAACCATTTCACTTCTCTAGTGTTTTAAAGTGGCTCGTGGCGAATTTG 3060
2220 GTTGTCTCTCTCGGAATAACACTCTCTTCCATGCTGTAGGGTGTGCTGCCAATGG 2279
3061 ATCCCTCAGGATATAGTGTGCTTTCGCTTTGCAAGGAGGGGAAATGATGCTTATGC 3120
2280 ATCCCTCAGGATATAGTGTGCTTTCGCTTTGCAAGGAGGGGAAATGATGCTTATGC 2339
3121 AATACTCTTGTAGTCTTTGCAACATGTTAAACGATGAGTTAGCAACATGCTTTTCAAGGAGA 3180
2340 AATACACTTGTAGTCTTTGCAACATGTTAAACGATGAGTTAGCAACATGCTTTTCAAGGAGA 2399
3181 GAAAAGCACCGTGTGCTGCGATTTGGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 3240
2400 GAAAAGCACCGTGTGCTGCGATTTGGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2459
3241 AAGCAACAGAGGCTGTGACATGAGTTGGACGAACCACTGAATTTCCGCAATTGCGAGAGAT 3300
2460 AAGCAACAGAGGCTGTGACATGAGTTGGACGAACCACTGAATTTCCGCAATTGCGAGAGAT 2519
3301 -ATTGTATTTAAGTGGCTGTGCTCGATACAGCAAAACGCCA--TTTGACCAATTCACCAATTT 3357

Db 2520 AATTGTAATTAAAGTGCCCTAGCTCGATACAGCAAAAGCCCAATTTTGAACCATTCACCAACATT 2579
Qy 3358 GGTGTGCACTT--CCAAGCTTCAAGCTGCGCAGCACTCAGGGCGCAAGGGCTGCTAAA 3415
Db 2580 GGTGTGCACTTCCAAAGCTTCAAGCTGCGCAGCACTCAGGGCGCAAGGGCTGCTAAA 2639
Qy 3416 GGAAGCGGAAACAGTGAAGAACAGTCCGAGAAAGCGTGTGACCCCGGATGAATGTCA 3475
Db 2640 GGAAGCGGAAACAGTGAAGAACAGTCCGAGAAAGCGTGTGACCCCGGATGAATGTCA 2699
Qy 3476 GCTACTGGGCTATCTGGCAAGAGGAAACGCAAGCGCAAGAGAAAGCAGGTAGCTTGCA 3535
Db 2700 GCTACTGGGCTATCTGGCAAGAGGAAACGCAAGCGCAAGAGAAAGCAGGTAGCTTGCA 2759
Qy 3536 GTGGGCTTACATGGGATAGTACGCTGGGCGGTTTATGACACAGCAAGCAAGCGGAAT 3595
Db 2760 GTGGGCTTACATGGGATAGTACGCTGGGCGGTTTATGACACAGCAAGCAAGCGGAAT 2819
Qy 3596 TGCCAGCTGGGGCGCCCTCTGTAAGGTTGGAAAGCCCTGCAAAAGTAACTGGATGGCTT 3655
Db 2820 TGCCAGCTGGGGCGCCCTCTGTAAGGTTGGAAAGCCCTGCAAAAGTAACTGGATGGCTT 2879
Qy 3656 TCTTGCCGCCAAGGATCTGATGGCGCAGGGATCAAGATCTGATCAAGAGACAGGATGAG 3715
Db 2880 TCTTGCCGCCAAGGATCTGATGGCGCAGGGATCAAGATCTGATCAAGAGACAGGATGAG 2939
Qy 3716 GATGTTTTCGATGATGAAACAGATGATGACAGCGAGGTTCTCCGGCGCTGGGGTG 3775
Db 2940 GATGTTTTCGATGATGAAACAGATGATGACAGCGAGGTTCTCCGGCGCTGGGGTG 2999
Qy 3776 AGAGGCTATTGGGCTATGACTGGGCACAAAGACAACTGGGCTGCTGATGCGCGCGTGT 3835
Db 3000 AGAGGCTATTGGGCTATGACTGGGCACAAAGACAACTGGGCTGCTGATGCGCGCGTGT 3059
Qy 3836 TCCGGCTGTACGCGCAGGGCGCGGTTCTTTTGTCAAGACCGACCTGTCGGTGCCC 3895
Db 3060 TCCGGCTGTACGCGCAGGGCGCGGTTCTTTTGTCAAGACCGACCTGTCGGTGCCC 3119
Qy 3896 TGAATGAATCGAGACAGGAGCGCGGCTATCGTGCTGCGCAGCGAGGCTGTCCTT 3955
Db 3120 TGAATGAATCGAGACAGGAGCGCGGCTATCGTGCTGCGCAGCGAGGCTGTCCTT 3179
Qy 3956 GCGCAGCTGTCTGACGCTGTCTACTGAAGCGGAGAGGACTGCTGCTATTGGCGGCAAG 4015
Db 3180 GCGCAGCTGTCTGACGCTGTCTACTGAAGCGGAGAGGACTGCTGCTATTGGCGGCAAG 3239
Qy 4016 TGCGGGGCGAGGATCTCCTGTATCTCACTTGTCTCTGCGAGAAAGTATCCATCATGG 4075
Db 3240 TGCGGGGCGAGGATCTCCTGTATCTCACTTGTCTCTGCGAGAAAGTATCCATCATGG 3299
Qy 4076 CTGATGCAATCGGGGCTGATAGCTGTTGATCGGCTACTGCGGCTGCGCAGCAACCAAG 4135
Db 3300 CTGATGCAATCGGGGCTGATAGCTGTTGATCGGCTACTGCGGCTGCGCAGCAACCAAG 3359
Qy 4136 CGAAACATCGGATCGAGCGACGATCTCTGATGGAAGCGGCTGTTGCGATCAGGATG 4195
Db 3360 CGAAACATCGGATCGAGCGACGATCTCTGATGGAAGCGGCTGTTGCGATCAGGATG 3419
Qy 4196 ATCTGGAGAAAGAGATCAGGGGCTGCGCGCAGCGCAACTGTTGCGGAGGCTCAAGGCGC 4255
Db 3420 ATCTGGAGAAAGAGATCAGGGGCTGCGCGCAGCGCAACTGTTGCGGAGGCTCAAGGCGC 3479
Qy 4256 GCATGCCCGAGCGGAGGATCTGCTGTGACCCATGCGGATGCTGCTGCGGATATCA 4315
Db 3480 GCATGCCCGAGCGGAGGATCTGCTGTGACCCATGCGGATGCTGCTGCGGATATCA 3539
Qy 4316 TGCTGGAATAGCGGCTTTCTGATTCATGACTGTGGCGGCTGGGCTGGCGGAC 4375
Db 3540 TGCTGGAATAGCGGCTTTCTGATTCATGACTGTGGCGGCTGGGCTGGCGGAC 3599
Qy 4376 GCTATCAGGACATAGCTGTGGCTACCGCTGATATTGCTGAAGAGCTTGGCGCGGAATGGG 4435

Db 3600 GCTATCAGGACATAGCGTTGGCTACCGGTGATATTGCTGAAGAGCTTTGGCGCGAATGGG 3659
Qy 4436 CTGACCGCTTCCCTCGTCTTTTACGGTATCGCGCTCCCGATTTCGACGCGCATCGCTTCT 4495
Db 3660 CTGACCGCTTCCCTCGTCTTTTACGGTATCGCGCTCCCGATTTCGACGCGCATCGCTTCT 3719
Qy 4496 ATGCGCTTCTTGACGAGTCTTCTTGAGCGGAGCTCTCGGGGTTTCGATA 4542
Db 3720 ATGCGCTTCTTGACGAGTCTTCTTGAGCGGAGCTCTCGGGGTTTCGATA 3766

RESULT 3
US-09-897-511A-9
; Sequence 9, Application US/09897511A
; Publication No. US20030092882A1
; GENERAL INFORMATION:
; APPLICANT: Bremel, Robert
; APPLICANT: Miller, Linda
; APPLICANT: Black, Gregory
; TITLE OF INVENTION: Host Cells Containing Multiple Integrating Vectors
; FILE REFERENCE: GALA-06416
; CURRENT APPLICATION NUMBER: US/09/897,511A
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 60/215,925
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 5130
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-09-897-511A-9

Query Match 30.7%; Score 1588.6; DB 3; Length 5130;
Best Local Similarity 66.1%; Pred. No. 0;
Matches 3007; Conservative 0; Mismatches 754; Indels 786; Gaps 18;

Qy 1 TTTGAAAGACCCACCCGTAGGTGGCAAGCTTAGCTTAAGCTAAAGCCACTTTGCAAGGCAT 60
Db 1 TTTGAAAGACCCACCCGTAGGTGGCAAGCTTAGCTTAAGCTAAAGCCACTTTGCAAGGCAT 60
Qy 61 GGAATAATACATACTAGTATGAGAAAGTTTCAAGTCAAGGTCAAGAAAGAAAGCAAGC 120
Db 61 GGAATAATACATACTAGTATGAGAAAGTTTCAAGTCAAGGTCAAGAAAGAAAGCAAGC 120
Qy 121 TGAATACCAACAGGATATCTGTGTAAGCGGTTTCTGCCCCGGCTCAGGGCCCAAGAAC 180
Db 121 TGAATACCAACAGGATATCTGTGTAAGCGGTTTCTGCCCCGGCTCAGGGCCCAAGAAC 180
Qy 181 GATGAGACAGCTGATGATGGGCAACAGGATATCTGTGTAAGCGGTTTCTGCCCCGG 240
Db 181 GATGAGACAGCTGATGATGGGCAACAGGATATCTGTGTAAGCGGTTTCTGCCCCGG 240
Qy 241 CTGGGGCCCAAGAACAGATGCTCCCAAGTCCGAGTCCAGCCCTCAGCAGTTTCTAGTGAA 300
Db 241 CTGGGGCCCAAGAACAGATGCTCCCAAGTCCGAGTCCAGCCCTCAGCAGTTTCTAGTGAA 300
Qy 301 TCATCAGATGTTTCCAGGGTGCACCAGGACCTGAAATGACCCCTGTACCTTATTTGAAC 360
Db 301 TCATCAGATGTTTCCAGGGTGCACCAGGACCTGAAATGACCCCTGTACCTTATTTGAAC 360
Qy 361 TAACCAATCAGTTGCTGCTTCTGCTTCTGTTGCGCGCTTCGGCTCTCGAGCTCAATAAA 420
Db 361 TAACCAATCAGTTGCTGCTTCTGCTTCTGTTGCGCGCTTCGGCTCTCGAGCTCAATAAA 420
Qy 421 AGAGCCCAACCCCTCACTCGCGCGCAGTCTTCGATAGACTGCGTCCCGGGTAC 480
Db 421 AGAGCCCAACCCCTCACTCGCGCGCAGTCTTCGATAGACTGCGTCCCGGGTAC 480
Qy 481 CCGTATTTCCCAATAAAGCCTCTTGTCTTTTGCATCCGAATCGTGTCTGCTGTTCTTTG 540

Db 481 CCGTATTCCTCCATTAAGCCTCTTGCTGTTTGGATCCGAATCGTGGTCTCGCTGTTCTCTTG 540
Qy 541 GGAGGGTCTCTCTGAGTGAATGACTACCCACGACGGGGCTCTTTCATTTTGGGGCTCGT 600
Db 541 GGAGGGTCTCTCTGAGTGAATGACTACCCACGACGGGGCTCTTTCATTTTGGGGCTCGT 600
Qy 601 CCGGGATTTGAGACCCCTCCAGGGACACCGACCCACCCAGGGAGGTAAAGTGGCC 660
Db 601 CCGGGATTTGAGACCCCTCCAGGGACACCGACCCACCCAGGGAGGTAAAGTGGCC 660
Qy 661 AGCAACTTATCTGTCTGCTCGGATGCTAGTGTCTATGTTGATGTTATGCGCTGCG 720
Db 661 AGCAACTTATCTGTCTGCTCGGATGCTAGTGTCTATGTTGATGTTATGCGCTGCG 720
Qy 721 TCTGTACTAGTAACTAGTCTGTATCTGCGGACCCGTTGGAACTGACGAGTT 780
Db 721 TCTGTACTAGTAACTAGTCTGTATCTGCGGACCCGTTGGAACTGACGAGTT 780
Qy 781 CTGACACCCCGCCGACCTGGGAGACGTCACAGGACCTTTGGGGCCGCTTTTGTGG 840
Db 781 CTGACACCCCGCCGACCTGGGAGACGTCACAGGACCTTTGGGGCCGCTTTTGTGG 840
Qy 841 CCGGACCTGAGGAAGGAGTGCATGTGAATCCGACCCCGTCAGGATATGTTCTGTT 900
Db 841 CCGGACCTGAGGAAGGAGTGCATGTGAATCCGACCCCGTCAGGATATGTTCTGTT 900
Qy 901 AGGAGACGAGAACCTTAAACAGTTTCCCGCTCCGCTCTGAATTTTGTCTTTCGTTTGGAA 960
Db 901 AGGAGACGAGAACCTTAAACAGTTTCCCGCTCCGCTCTGAATTTTGTCTTTCGTTTGGAA 960
Qy 961 CCGAAGCCGCGCTGTTGCTGTCGACGCAAGCTTGGGCTGCGAGTCTGAGGAA 1020
Db 961 CCGAAGCCGCGCTGTTGCTGTCGACGCAAGCTTGGGCTGCGAGTCTGAGGAA 1020
Qy 1021 TCAATTGCGCACGAGTAATCGGTGCTGCGCTCTTTAGGACATATGAAGTATGCACAGT 1080
Db 1021 CC-----CTGCACCCGAAACATGGAGAACCAACATCAGGATTCCT 1059
Qy 1081 GGGATGACTTTTCTGATCAGCAAGAGGACACTGACAGCTGTCACAGAGTCTGTGAAGTTG 1140
Db 1060 AGGACCCCTGCTGTTTACAGGGGGGTTTTTCTTGTGTGACAGATCTCTCACATACC 1119
Qy 1141 ATGCTCGCTCAGTACAGCTTTGCTTCTCCCAATCTTAAATGAGGCGCAACTCTTCAAG 1200
Db 1120 ACAGAGTACAGTCTGCTGCTGCTCTCTCAATTTCTAGGGGAGCACCCAGTGTCC 1179
Qy 1201 AGAGGATGAGTCTTATAAAGTCACTGATGACCCCTTATCTCAATGTTGTTGTAGTTC 1260
Db 1180 TGGCCAAATTCGAGTCCCACTCCCAATCACTCACCACCTCTTGTCTCTCAATTTG 1239
Qy 1261 TCGTGCCCATCATTTGGCATAGTGGCAGCTCAGCTCTCTGAATGGGAAACGAAGAAATGCA 1320
Db 1240 TCCTGGCTATC-----GCTGGATGTTCTGCGGGTTTTATC 1276
Qy 1321 CGGTTGGCTCAGTTAATGACAGATATCTCCAACTCCGGAAGGCAAGAAATGCGAGTG 1380
Db 1277 ATATTCTCTTATCTCTGCTATGCTCATCTCTTGTGTTGTTCTCTGGAATAACAA 1336
Qy 1381 AAGATGAATTCAGATTTTCGAGAGCTGTGAGAACGATGACCAATGGAAGACAGAA 1440
Db 1337 GGTATTTGGCCGTTTGTCTCTACTTCCAGGAA---CATCAATACAGCACCGGACCA 1393
Qy 1441 TCCAGTATCTTTCAGATAATGAAGCAATCTCTCTAGATGCTTAAGAAATTTCCAAATTTCA 1500
Db 1394 TGGAAAGCTTCAGATTCCTGCTCAAGGAACCTCTATGTTTCCCTCTT-----GTTGCT 1448
Qy 1501 GCATAACAACTGATCAAGATTTAATGATGTTCTTTTCCAGCTAAATTCCTTACTTCTCT 1560
Db 1449 GTACAAACCTTCGAGCGGAACTGCACCTGTATTTCCCATCCATCTCTGGGCTTTG 1508
Qy 1561 CCATCCAGGACATGAGATATCATAGGGGATATCTCCAACTCATTTAGTGTCTGAACA 1620
Db 1509 CAAGATTCCTATGGGAGTGGGCCCTCAGTCCGTTTCTCTCTG----- 1549

Qy 1621 CCACAGTACTTGAATTTGACGTTTCAAGTATTTGAAACACATGAATGGCAGAGTCCAAGAGAAATG 1680
Db 1550 -CTCAGTTTACTAGTGCAATTTGTTCAAGTGTTC---GTAGGGCTTTTCCCACTGTTTG 1605
Qy 1681 CATTTAAACACAGAGGAGATGCGTAATTTAGAGGAGCGTATATACAAATGCAATCAGCAG 1740
Db 1606 GCTTTCAAGTTATATG-GATGATGTGTTATTTGGGGCCCAAGTCTGTACAAATCTTTAGTTC 1664
Qy 1741 AAATTAAGTCTCTAGATGAAACCAAGTATATTTTGGAAACAGGAATAAAGGGGAAATGA 1800
Db 1665 CTTTTTACCTCTATTAACCAATTTTCTTTGCTCTTTGGGTATACATTTAAACCCCTAATAA 1724
Qy 1801 AACTGTTGAATAATATCACTAATGATCTGAGGCTGAAGGATTTGGGAACATTTCTCAGACAT 1860
Db 1725 AAC-----CAACGCT 1734
Qy 1861 TGAATAATATACATTTTACTCCAGGTGCCAGAAAGTGTCTGCTGACTGGGAAATGGACCA 1920
Db 1735 TGGGGCTACTCCCTTAACTTCATGGGATATGTAATTTGGATGTTGGG----- 1781
Qy 1921 ACATCTGGGCTCCAAACATGACCATCGGGGCTGTGAACAGCAGAGGTGAATTCACAGGCA 1980
Db 1782 ----- 1781
Qy 1981 CCTACATCACAGCGTAAACAGCCACATCAATAGATCAAAAGTCAACACTGATGGGA 2040
Db 1782 -----TACTTTACCGCAAGACATATTTGCTACTAAAAATCAAGCAATGTTTTCG- 1829
Qy 2041 CACAAACACATCAACAGAGGACCCAGGCCACCTTTGGCTTCCAGTCAATTTGGAAGT 2100
Db 1830 -----AAAACCTGCTGTAAATAGACCTTATTTGTTGAAAGTATGTCAGAGACTTTGTTGGGT 1884
Qy 2101 TTTCAGAGTCCACCACTGCTTTACCGGCCAGTCTTCATAGACAGGAATGGGAAGGAGG 2160
Db 1885 CTTTTGGGCTTTGCTGCCCTTTTACAAATGSGCTATCTCTGCTTAAATGCTTTTATAT 1944
Qy 2161 TCTTGAAGACATGTTGCTGCTGCGGTCAAGTGTAAATGACATTTGTTGATGACTGGAAAG 2220
Db 1945 GCATGTATACATCT----- 1959
Qy 2221 CTACAGGGTGGGATCAACATCTTCACTGCTGCGCACACAGAAAGAGTGAAGTGAAGT 2280
Db 1960 -----AAGCAGGCTTTTCACTTTCTCGCAACTTACAAAGCCCTTTCTGTGTA 2005
Qy 2281 ACCAAGTCTCTCTGGACTCCAGGTGAAGAGATAGAGGCCCTCTCTGCAAAAATGG 2340
Db 2006 AACATATCTGAACCTTTTACCCCGTTGCCCGGCA----- 2039
Qy 2341 TATACAGGCTTTTCCAGGTCTAATAGGTACTTCCAGGTCTTAAAGTGTGATCGGGGGATCT 2400
Db 2040 ----- 2039
Qy 2401 CTGGTTTACCTGGAGTTTCGAGGATTCACAGCAACCAATGGGGAGAACCGGGAGCCAGGAC 2460
Db 2040 ----- 2039
Qy 2461 TTAATGACAAAAGGCCAGAGGGGAGAAAAGGGAGTGGNAGCATGCMAAGACAAATCTA 2520
Db 2040 -----ACGGTCAG 2047
Qy 2521 ATACAGTCCGACTGTTGGTGGCAGCGGCCCTTCAAGAGGAGAGTGGAGATTTTTCACG 2580
Db 2048 GTCTCTCCAAAGTGTGTGTGACGCAACCCCACTGGATGGGCTTGGCTATCGGCCATA 2107
Qy 2581 AAGGCCAGTGGGTACCGTGTGTGACGACCGCTGGGAACTGCGGTGGAGGACTGGTCTGTCT 2640
Db 2108 GCGCATGCGCGGACCTTTTGTGGCTCTCTGCGG----- 2141
Qy 2641 GCAGGAGCTTTGGGATCAAAAGGTGTTCAAAGTGTGCATAGCGAGCTTATTTTGGAAAAG 2700
Db 2142 -----ATCCATACTCGGAACTCTTAGCAGCTTGT----- 2171

ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic
US-10-397-079-9

Query Match 30.7%; Score 1588.6; DB 6; Length 5130;
Best Local Similarity 66.1%; Pred. No. 0;
Matches 3007; Conservative 0; Mismatches 754; Indels 786; Gaps 18;

| | | | |
|----|-----|---|-----|
| Qy | 1 | TTTGAAGACCCACCCCTAGTGGTGAAGCTAGCTTAAAGTAAAGCCACTTTGCAAGGCAT | 60 |
| Db | 1 | TTTGAAGACCCACCCCTAGTGGTGAAGCTAGCTTAAAGTAAAGCCACTTTGCAAGGCAT | 60 |
| Qy | 61 | GGAAAAATACATACTAGATAGATAAGAAAAGTTTCAAGTCAAGGTCAAGAAACAAGAAACAGC | 120 |
| Db | 61 | GGAAAAATACATACTAGATAGATAAGAAAAGTTTCAAGTCAAGGTCAAGAAACAAGAAACAGC | 120 |
| Qy | 121 | TGAATACCAAAACAGGATATCTGTGTAGAGCGTTCCTGCCCGGCTCAGGGCCAAAGAAC | 180 |
| Db | 121 | TGAATACCAAAACAGGATATCTGTGTAGAGCGTTCCTGCCCGGCTCAGGGCCAAAGAAC | 180 |
| Qy | 181 | GATGAGACAGCTAGTGTAGTGGCCAAACAGGATATCTGTGTAGAGCGTTCCTGCCCGG | 240 |
| Db | 181 | GATGAGACAGCTAGTGTAGTGGCCAAACAGGATATCTGTGTAGAGCGTTCCTGCCCGG | 240 |
| Qy | 241 | CTGGGGCCAAAGAACAGATAGTTCCTCAGATGCGGTCCAGCCCTCAGCAGTTTCTAGTGAA | 300 |
| Db | 241 | CTGGGGCCAAAGAACAGATAGTTCCTCAGATGCGGTCCAGCCCTCAGCAGTTTCTAGTGAA | 300 |
| Qy | 301 | TCATCAGATGTTCCAGGGTCCCAAGGACCTGAAATGACCTGTACCTTATTGAAAC | 360 |
| Db | 301 | TCATCAGATGTTCCAGGGTCCCAAGGACCTGAAATGACCTGTACCTTATTGAAAC | 360 |
| Qy | 361 | TAAACAATCAGTTCGCTTCGTCTGTTCGCGGCTTCGCGCTCCGAGCTCAATAAA | 420 |
| Db | 361 | TAAACAATCAGTTCGCTTCGTCTGTTCGCGGCTTCGCGCTCCGAGCTCAATAAA | 420 |
| Qy | 421 | AGAGCCCAACCCCTCACTCGGGCCGACGCTTCGATAGAGTGCCTCGCCGGGTAC | 480 |
| Db | 421 | AGAGCCCAACCCCTCACTCGGGCCGACGCTTCGATAGAGTGCCTCGCCGGGTAC | 480 |
| Qy | 481 | CGGTATTCCTCAATAAGCTTCGTGTTCGATCCGATCGTGTCTCGCTGTCCTTG | 540 |
| Db | 481 | CGGTATTCCTCAATAAGCTTCGTGTTCGATCCGATCGTGTCTCGCTGTCCTTG | 540 |
| Qy | 541 | GGAGGCTCTCTCTGAGTGATTGACTACCCAGCGGGGCTCTTCATTTGGGGCTCGT | 600 |
| Db | 541 | GGAGGCTCTCTCTGAGTGATTGACTACCCAGCGGGGCTCTTCATTTGGGGCTCGT | 600 |
| Qy | 601 | CCGGATTGGAGACCCCTGCCAGGGACCAACCCAGCCAGCGGAGGTAAAGTGGCC | 660 |
| Db | 601 | CCGGATTGGAGACCCCTGCCAGGGACCAACCCAGCCAGCGGAGGTAAAGTGGCC | 660 |
| Qy | 661 | AGCAACTATCTGTCTGTGGATGCTAGTGTCTATGTTGATGTTATGCGGCTGCG | 720 |
| Db | 661 | AGCAACTATCTGTCTGTGGATGCTAGTGTCTATGTTGATGTTATGCGGCTGCG | 720 |
| Qy | 721 | TCTGTACTAGTTAGCTAGCTAGCTATCTGGCGGACCCGCTGGTGGAACTGACGAGTT | 780 |
| Db | 721 | TCTGTACTAGTTAGCTAGCTAGCTATCTGGCGGACCCGCTGGTGGAACTGACGAGTT | 780 |
| Qy | 781 | CTGAACACCCCGCCCAACCCCTGGGAGACGCTCCAGGGACCTTTGGGGCCGCTTTTGTGG | 840 |
| Db | 781 | CTGAACACCCCGCCCAACCCCTGGGAGACGCTCCAGGGACCTTTGGGGCCGCTTTTGTGG | 840 |
| Qy | 841 | CCCGACTTGGAGAGGGAGTGATGTGAAATCCGACCCCGCTCAGGATATGTTGTTCTGGT | 900 |
| Db | 841 | CCCGACTTGGAGAGGGAGTGATGTGAAATCCGACCCCGCTCAGGATATGTTGTTCTGGT | 900 |
| Qy | 901 | AGGAGACGAGACCTTAAACAGTTTCCCGCTCCGCTGAAATTTTTCGTTTGGAA | 960 |
| Db | 901 | AGGAGACGAGACCTTAAACAGTTTCCCGCTCCGCTGAAATTTTTCGTTTGGAA | 960 |

| | | | |
|----|------|--|------|
| Qy | 961 | CCGAAGCCGCGGCTTGTCTCTGCGACCCCAAGCTTGGGCTGCGAGTCTCTAGAGA | 1020 |
| Db | 961 | CCGAAGCCGCGGCTTGTCTCTGCGACCCCAAGCTTGGGCTGCGAGTCTCTAGAGA | 1020 |
| Qy | 1021 | TCAATTCGGCACGAGTAAATCGGTCTTTCAGGACATATGAAGTATGGCACAGT | 1080 |
| Db | 1021 | CC-----CTGCACCGAANATGGAGAACAAATCAGGATTCCT | 1059 |
| Qy | 1081 | GGGATGACTTTCTGATCAGCAAGAGGACACTGACAGCTGTACAGAGTCTGTGAAGTTGG | 1140 |
| Db | 1060 | AGGACCCCTGCTCGTGTACAGGGGGGTTTTCTTGTGTGACAGAACTCTCACAATACC | 1119 |
| Qy | 1141 | ATGCTCGCTCAGTGACAGCTTTGCTCTCCCATCTCTAAATATGGCCCAACTCTCAAG | 1200 |
| Db | 1120 | ACAGAGTCTAGACTGCTGGTGGACTTCTCTCAATTTCTAGGGGGAGCACCCACGTGCC | 1179 |
| Qy | 1201 | AGAGATGAAGTCTTATAAAATCGCAGTATCACCTTTTATCTATCTGTTGTAGTTTC | 1260 |
| Db | 1180 | TGGCCAAATTCGCGAGTCCCAACCTCAATCACTCACCACCTTGTCTCTCAATTTG | 1239 |
| Qy | 1261 | TCGTGCCCATCAATTGGCAGTCTAGTGGCAGCTCAGCTCTGAAATGGGAAACGAAGATTGA | 1320 |
| Db | 1240 | TCCTGGCTATC-----GCTGGATGTGTCTCGCGGCTTTATC | 1276 |
| Qy | 1321 | CGGTTGGCTCAGTTAATGACAGATATATCTCAAGTCCGGAAGGCAAGAAATGGCAGTG | 1380 |
| Db | 1277 | ATAATCTCTCTCATCTGCTATGCTCTATCTTCTTGTGTGTTCTCTTGGACTACCAA | 1336 |
| Qy | 1381 | AAGATGAATCAGATTTTCGAGAACTGTGAGGACGATGAGCAACATGGAAGACAGAA | 1440 |
| Db | 1337 | GGTATGTTGCCGCTTGTCTCTCTCTCAGGAA---CATCAACTACCAGCAGCGGACCA | 1393 |
| Qy | 1441 | TCCAGTATCTTTTCAGATAATGAAGCCAATCTCTAGATGCTAAGAAATTTCCAAAATTTCA | 1500 |
| Db | 1394 | TGCAAGACCTCGACGATCTCTCTCAAGGAACCTATGTTTCCCTCTT-----GTGCT | 1448 |
| Qy | 1501 | GCATAACAACTGATCAAGATTTAATGATGTTTCTTCCAGTAAATTCCTTACTTCTCT | 1560 |
| Db | 1449 | GTACAAAACCTTCGACGGAACCTGCACTTGATATTCCTCATCCATCATCTGGGCTTCG | 1508 |
| Qy | 1561 | CCATCCAGGACATGAGAAATATCATAGGGGATATCTCCAAGTCTATAGTCTGTAACA | 1620 |
| Db | 1509 | CAAGATTCCTATGGAGTGGGCTCAGTCCGTTCTCTCTCG-----1549 | |
| Qy | 1621 | CCACAGTACTTGAATTTGAGTTCAGTATTTGAAACAATGAAATGGCAGAGTCCAAAGAAATG | 1680 |
| Db | 1550 | -CTCAGTTACTAGTGCCATTTGTTTCTAGTGGTCT---GTAGGGCTTTCCCCCACTGTTG | 1605 |
| Qy | 1681 | CATTAAACAAAGAGAGATGCGTAAATTTAGAGGAGCGTATATACAAATGATCAGCAG | 1740 |
| Db | 1606 | GCTTTCAGTTATATG-GATGATGTTGTTTGGGGGCCAAGTCTGTACAACTCTTGAGTC | 1664 |
| Qy | 1741 | AAATTAAGTCTCTAGATGAAACAAAGTATATTTGGAAACAGAAATAAAGGGGAAATGA | 1800 |
| Db | 1665 | CTTTTTCACCTCTATTACCAATTTTCTTGTCTTGGGTATACATTTTAAACCCCTAATAA | 1724 |
| Qy | 1801 | AACGTTGAATAATCACTAATGATCTGAGGCTGAAGGATTTGGGAACATCTCTCAGACAT | 1860 |
| Db | 1725 | AAC-----CAAAGT | 1734 |
| Qy | 1861 | TGAAAAATATCACTTTTACTCCAAGTGCAGAAAGTGTCTGCTGATCGGGAATGGAACA | 1920 |
| Db | 1735 | TGGGGCTACTCCCTTAACTTCAATGGGATATGTAATTTGGATGTTGGG-----1781 | |
| Qy | 1921 | ACGATCTGGGCTCCAAATGACCATCGGGGCTGTGAACAGCAGAGGTTGAATTTACAGGCA | 1980 |
| Db | 1782 | -----1781 | |
| Qy | 1981 | CCTACATCACAGCGGTAAAGCCACATCAATGAGATCAAAAGAGTTCACCACTGCAATGGGA | 2040 |
| Db | 1782 | -----TACTTTACCCCAAGAACATATTTGTAATAAATAACAAGCAATTTTTCG-1829 | |
| Qy | 2041 | CACAAAACACCATCAACAAAGAGGAGCCAGCCACCTTTGGCTTCCACCGTCAATTTGGAAGT | 2100 |

[illegible]

| | | | |
|----|------|---|------|
| Db | 2340 | AAATACACTTGTAGTCTTTGGCAACAATGGTAAACGATGAGTTTAGCAACATGCGCTTACAGAGAGA | 2399 |
| Qy | 3181 | GA AAAAGCACCGTGCATCGCCGATTTGGTGGAAGTAAAGGTGGTACGATCGTGCCTTATTAGG | 3240 |
| Db | 2400 | GA AAAAGCACCGTGCATCGCCGATTTGGTGGAAGTAAAGGTGGTACGATCGTGCCTTATTAGG | 2459 |
| Qy | 3241 | AAGGC AAACAGACGGGTCTGACATGGAATTGGACGAAACCATGGAATTCGCGAATTCGACAGAT | 3300 |
| Db | 2460 | AAGGC AAACAGACAGGTCTGACATGGAATTGGACGAAACCATGGAATTCGCGAATTCGACAGAT | 2519 |
| Qy | 3301 | -ATTGTATTTAAAGTGGCTAGCTCGATACAGCAAAACGCCA--TTTGTACCAATTACACCAATT | 3357 |
| Db | 2520 | AA TTTGTATTTAAAGTGGCTAGCTCGATACAGCAAAACGCCAATTTTGTACCAATTACACCAATT | 2579 |
| Qy | 3358 | GGTGTGCACTCT--CCAAAGCTTCAACGTGCGCAAGCACTCAGGCGCAAGGGCTGCTAAA | 3415 |
| Db | 2580 | GGTGTGCACTCTCCAAAGCTTCAACGTGCGCAAGCACTCAGGCGCAAGGGCTGCTAAA | 2639 |
| Qy | 3416 | GGAAAGCGGAACAAGTAGAAAGCCAGTCCGCAGAAAAGGTTGTAACCCCGGATGAATGTCA | 3475 |
| Db | 2640 | GGAAAGCGGAACAAGTAGAAAGCCAGTCCGCAGAAAAGGTTGTAACCCCGGATGAATGTCA | 2699 |
| Qy | 3476 | GCTACTGGGCTATCTGGACAAGGGA AAAACGCAAGCGCAAGAGAGAAAGCAGGTAGCTTGCA | 3535 |
| Db | 2700 | GCTACTGGGCTATCTGGACAAGGGA AAAACGCAAGCGCAAGAGAGAAAGCAGGTAGCTTGCA | 2759 |
| Qy | 3536 | GTGGGCTTACATGGCGATAGCTAGACTGGGCGGTTTTATGGACGCAAGCGAACCCGGAAT | 3595 |
| Db | 2760 | GTGGGCTTACATGGCGATAGCTAGACTGGGCGGTTTTATGGACGCAAGCGAACCCGGAAT | 2819 |
| Qy | 3596 | TGCCAGCTGGGGCGCCCTCTGGTAAAGGTTGGGAAGCCCTGCAAAAGTAAACTGGAATGGCTT | 3655 |
| Db | 2820 | TGCCAGCTGGGGCGCCCTCTGGTAAAGGTTGGGAAGCCCTGCAAAAGTAAACTGGAATGGCTT | 2879 |
| Qy | 3656 | TC TTTCCGCCCAAGGATCTGATGCGCAGGGGATCAAGATCTGATCAAGAGACAGGATGAG | 3715 |
| Db | 2880 | TC TTTCCGCCCAAGGATCTGATGCGCAGGGGATCAAGATCTGATCAAGAGACAGGATGAG | 2939 |
| Qy | 3716 | GATCGTTTCCGCTATGATTAACAAGATGGATTGCAACGACAGTCTCTCCGCGCGCTTGGGTGG | 3775 |
| Db | 2940 | GATCGTTTCCGCTATGATTAACAAGATGGATTGCAACGACAGTCTCTCCGCGCGCTTGGGTGG | 2999 |
| Qy | 3776 | AGAGGCTATTCCGCTATGACTGGGCAACAACAGACAATCGGCTGCTCTGATGCGCGCGTGT | 3835 |
| Db | 3000 | AGAGGCTATTCCGCTATGACTGGGCAACAACAGACAATCGGCTGCTCTGATGCGCGCGTGT | 3059 |
| Qy | 3836 | TCGGGCTGTACGCGCAGGGGCGCCGGTCTCTTTTGTCAAGACCGACCTGTCGGGTGCC | 3895 |
| Db | 3060 | TCGGGCTGTACGCGCAGGGGCGCCGGTCTCTTTTGTCAAGACCGACCTGTCGGGTGCC | 3119 |
| Qy | 3896 | TGAATGAATGCAAGGACGAGGCGCGGCTATCGTGGCTGGCCACGACGGGGCTTCCCTT | 3955 |
| Db | 3120 | TGAATGAATGCAAGGACGAGGCGCGGCTATCGTGGCTGGCCACGACGGGGCTTCCCTT | 3179 |
| Qy | 3956 | GCGCAGCTGTCTCGAGTTGTCACTAAAGCGGGAAGGGGACTGGCTGCTATTGGGCGGAAG | 4015 |
| Db | 3180 | GCGCAGCTGTCTCGAGTTGTCACTAAAGCGGGAAGGGGACTGGCTGCTATTGGGCGGAAG | 3239 |
| Qy | 4016 | TGCCGGGGCAGGATCTCTGTGTCACTCACTTGTCTCTGCCGAGAAAGTATCCATCATGG | 4075 |
| Db | 3240 | TGCCGGGGCAGGATCTCTGTGTCACTCACTTGTCTCTGCCGAGAAAGTATCCATCATGG | 3299 |
| Qy | 4076 | CTGATGCAATGGCGGCGCTCATACGCTTGAATCCGCTACTCTCCCACTTCGACCAACAAG | 4135 |
| Db | 3300 | CTGATGCAATGGCGGCGCTCATACGCTTGAATCCGCTACTCTCCCACTTCGACCAACAAG | 3359 |
| Qy | 4136 | CGAAACATCGCATCGACGACGACGTA CTGGAATGGAAGCCGGTCTTGTGATCAGGATG | 4195 |
| Db | 3360 | CGAAACATCGCATCGACGACGACGTA CTGGAATGGAAGCCGGTCTTGTGATCAGGATG | 3419 |
| Qy | 4196 | ATCTGGAACGAAGACATCAGGGGCTCGCGCAGCGCAACTGTTTCGCGAGGCTCAAGGCGC | 4255 |
| Db | 3420 | ATCTGGAACGAAGACATCAGGGGCTCGCGCAGCGCAACTGTTTCGCGAGGCTCAAGGCGC | 3479 |

Db 1394 TCGAAGACCTGCACGATTCCTGCTCAAGAACCTCTAATGTTTCCCTCTT-----GTTGCT 1448
Qy 1501 GGAATACAACTGATCAAGATTTAATGATGTTCTTTTCCAGGTAAATTCCTTACTTTCCT 1560
Db 1449 GTACAAAACCTTCGGACGGAACCTGCATCTGTATTTCCCATCCATCCTCGGGCTTTGC 1508
Qy 1561 CCATCCAGGAACATGAGATATCATAGGGGATATCTCCAAGTCATTAGTAGGTCTGACAA 1620
Db 1509 CAAGATTCCTATGGAGTGGGCTCAGTCCGTTTCTCCTGG----- 1549
Qy 1621 CCACAGTACTTGAATTTGAGTTTCAGTATTTGAAACACTGAATGGCAGAGTCCAAAGAGATG 1680
Db 1550 -CTCAGTTTACTAGTGCCATTTGTTTCAGTGGTTC---GTAGGGCTTCCCCCACTGTTG 1605
Qy 1681 CATTTAAACAAAGAGGAGATGCGTAAATTAGAGGCGGTATATACAAATGCATCAGCAG 1740
Db 1606 GCTTTTCAGTTATATG-GATGATGTGGTATTTGGGGCCAAAGTCTGTACAAACATCTTGAGTC 1664
Qy 1741 AAATTAAGTCTTAGATGAAACAAAGTATATTTTGGACAGGAAATAAAGGGGAATGA 1800
Db 1665 CTTTTTACCTCTATTACCAATTTTCTTTGCTTTGGGTATACATTTTAAACCCCTAATAA 1724
Qy 1801 AACTGTTGAATAATATCACTAATGATCTGAGGCTGAAGGATTTGGGAACATTTCTCAGACAT 1860
Db 1725 AAC----- 1734
Qy 1861 TGAATAATATCACTTTACTCCAGGTGCCAGAAAGTCTCGCTGACTGGGAAATGGACCA 1920
Db 1735 TGGGCTACTCCCTTAACTTCATGGGATATGTAATTTGGATGTTGGG----- 1781
Qy 1921 ACGATCTGGCTCCAAACATGATCCATCGGGCTGTGMAACAGCAGAGGTGAAATTCACAGGCA 1980
Db 1782 ----- 1781
Qy 1981 CCTACATCACAGCGGTAAACAGCCACATCAATAGATCAAGAGTCAAGAGTCAACACTGCATGGGA 2040
Db 1782 -----TACTTTACCGCAAGAACATATTTGTAATAAATCAACGATGTTTCG- 1829
Qy 2041 CACAAAACCATCAACAGAGGACCCAGCCACCTTTGGCTTCAACCTCAATTTGGAAGT 2100
Db 1830 -----AAATGCTCTGAATAAGACCTATTGATTGGAAGTATGTACAGAGCTTTGGGT 1884
Qy 2101 TTTGAGTCCACACTGCTTTCAGGGCCAGTCTTCATAGACAGGAATGGGAGGAGG 2160
Db 1885 CTTTGGGCTTTGCTGCCCTTTTACAAATGTGGCTATCTGCTTAAATGCTTTATAT 1944
Qy 2161 TCCTGAAGACCATGTGGCTGCTGCGGTCAAGTGTTAATGACATTTGGTATGACTGGAAAG 2220
Db 1945 GCATGTATACATCT----- 1959
Qy 2221 CTACAGGGTCGGCATCAACATCTTCACTCGCTCGGCACACAGAGAGGTGAGTGTG 2280
Db 1960 -----AAGCAGGGCTTTTCACTTTCTGCCCAACTTACAAAGCCCTTCTGTGTA 2005
Qy 2281 ACCAAGGTCTCTCTGATCCAGGTGAAAGAGGATAGAGCCCTCTCTGACAAATGG 2340
Db 2006 AACAAATATCTGAACCTTTTACCCCGTGTCCCGCA----- 2039
Qy 2341 TATACAGGCTTTCCAGGTCTAATAGTACTCCAGGTCTTAAAGGTGATCGGGGGATCT 2400
Db 2040 ----- 2039
Qy 2401 CTGTTTACTGGAGTTTCAGGATTTCCAGGACCAATGGGGAAGACCGGAGCCAGGAC 2460
Db 2040 ----- 2039
Qy 2461 TTAATGGACAAAAGCCGAGAGGAGAGAAAAGGGAGTGGAGCATGCAAGACAACTTA 2520
Db 2040 -----ACGGTCAG 2047
Qy 2521 ATACAGTCCGACTGGTGGGTGACGGCCCTCAGAAAGGACAGAGTGGAGATTTTTCAG 2580

Db 2048 GTCTCTGCCAAGTGTGTTGCTGCACGAACCCCACTGATGGGGCTTGCTATATCGCCATA 2107
Qy 2581 AAGCCAGTGGGTACGGTGTGTGACGACCGCTGGGAACCTGCGTGGAGGACTGCTGCT 2640
Db 2108 GCGCATGCGCGGACCTTTGTGGCTCTCTGCGG----- 2141
Qy 2641 GCAGGAGCTTTGGGATACAAAGGTGTTCAAAAGTGTGCATAAGCGAGCTTATTTTGGAAAAG 2700
Db 2142 -----ATCCATCTGCGGAACCTCTAGCAGCTTGT----- 2171
Qy 2701 GTACGGTCCAAATATGCTGAATGAAGTATTTTGTTCGGGAAAGAGTCATTCATTTGAAG 2760
Db 2172 ----- 2171
Qy 2761 AGTCAGAAATTAGACAGTGGGGTGTGAGAGCCTGTTTCGACGACGAAAGATGCTGGGGTC 2820
Db 2172 -----TTTGCTCGCAGCGGCTGTGAGCGGAA 2197
Qy 2821 ACTTTGCACCTTACATAATGCATCATATTTTCATTCACATTTTTTAAACTGTTATAAAGTG 2880
Db 2198 ACTTATCGG----- 2207
Qy 2881 ATTTTTTCTTTGCTTCACTAAATCAGCTTAATAATATTTAAGAAACATAAGATTTT 2940
Db 2208 ----- 2207
Qy 2941 ATCCACAGAAAGGAAATTTTAAAAATCACTGGATAAACATATATAAATAGCTTCATATTT 3000
Db 2208 -----ACGCAACACTCT 2219
Qy 3001 GCTTCAAATACCAAGACCATTTTCAACTTCTTAGTGTGTTTAAAGTGGCTGCTGCGGAAATG 3060
Db 2220 GTTGTCTCTCTCGGAAATACACCTCTTTCATGGCTGTAGGGTGTGCTGCCAATGG 2279
Qy 3061 ATCCCTCAGGATATAGTAGTTTCGCTTTTCATAGGAGGGGAAATGTAGTCTTATGC 3120
Db 2280 ATCCCTCAGGATATAGTAGTTTTCGCTTTTGCATAGGAGGGGAAATGTAGTCTTATGC 2339
Qy 3121 AATACTCTTGTAGTCTTGCACATGTTAAACATGTTGGAAGTGTAGCAACATGCTTACAGGAGA 3180
Db 2340 AATACACTTGTAGTCTTGCACATGTTTAAACATGTTGGAAGTGTAGCAACATGCTTACAGGAGA 2399
Qy 3181 GAAAGCACCGTGCATGCGGATTTGGAAGTGAAGTGTGATCCGATCGTGCCTTATTAGG 3240
Db 2400 GAAAGCACCGTGCATGCGGATTTGGAAGTGAAGTGTGATCCGATCGTGCCTTATTAGG 2459
Qy 3241 AAGCAACAGACGGGTCTGACATGATTTGGAACCACTGAATTTCCGATTTCCAGAGAT 3300
Db 2460 AAGCAACAGACAGGTCTGACATGATTTGGAACCACTGAATTTCCGATTTCCAGAGAT 2519
Qy 3301 -ATTGTATTTAAGTGCCTAGCTCGATACAGCAAAAGGCA--TTTGACCATTTCCACCATTT 3357
Db 2520 AATTGTATTTAAGTGCCTAGCTCGATACAGCAAAAGGCAATTTTGTGACCATTTCCACCATTT 2579
Qy 3358 GGTGTGCACCT--CCAAGCTTTCAGCTGCCCAAGCACTCAGGGCGCAAGGGCTGCTAAA 3415
Db 2580 GGTGTGCACCTTCCAAAGCTTTCAGCTGCCCAAGCACTCAGGGCGCAAGGGCTGCTAAA 2639
Qy 3416 GGAAGCGGAAACACGTAGAAAGCCAGTCCGCAAGAACCGTGTGACCCCGGATGAATGCA 3475
Db 2640 GGAAGCGGAAACACGTAGAAAGCCAGTCCGCAAGAACCGTGTGACCCCGGATGAATGCA 2699
Qy 3476 GCTACTGGGCTATCTGGACAAAGGGAACCGCAAGCGGCAAGAGAGAGAGTAGTGTGCA 3535
Db 2700 GCTACTGGGCTATCTGGACAAAGGGAACCGCAAGCGGCAAGAGAGAGAGTAGTGTGCA 2759
Qy 3536 GTGGGCTTACATGGCGATAGCTAGCTGGGCGGTTTTTATGACAGCAAGCAACCGGAAT 3595
Db 2760 GTGGGCTTACATGGCGATAGCTAGCTGGGCGGTTTTTATGACAGCAAGCAACCGGAAT 2819
Qy 3596 TGCCAGCTGGGGCGCCCTCTGGTAAAGGTGGGAAGCCCTGCAAAAGTAACTGGATGGCTT 3655
Db 2820 TGCCAGCTGGGGCGCCCTCTGGTAAAGGTGGGAAGCCCTGCAAAAGTAACTGGATGGCTT 2879

841 CCCGACCTGAGGAGGAGTGAATGTGGAATCCGACCCCGTCAGGATATGTGGTCTCGGT 900
Db
841 CCCGACCTGAGGAGGAGTGAATGTGGAATCCGACCCCGTCAGGATATGTGGTCTCGGT 900
Qy
901 AGGAGACAGAACTTAAACAGTTCCCGCCTCCGTCGAAATTTTGTCTTCGGTTGGAA 960
Db
901 AGGAGACAGAACTTAAACAGTTCCCGCCTCCGTCGAAATTTTGTCTTCGGTTGGAA 960
Qy
961 CCGAAGCCGCGGTCTGTCTGCTGCAGCAGAGCTTGGGCTGCAGGTCTAGAGGA 1020
Db
961 CCGAAGCCGCGGTCTGTCTGCTGCAGCAGAGCTTGGGCTGCAGGTCTAGAGGA 1020
Qy
1021 TCMAATTCGGCACGAGTAATCGGTGCTGCGCTCTTATGAGACATATGAAGTATGCACAGT 1080
Db
1021 CC-----CTGCACCGAACATGGAGAACACACATCAGAAATTCCT 1059
Qy
1081 GGGATGACTTTCTGATCAGCAAGAGGACACTGACAGCTGTACAGAGTCTGTGAAGTTCG 1140
Db
1060 AGGACCCCTGCTCGTGTACAGGGGGGTTTTCTTGTGTGACAAAGATCTCTCAATACC 1119
Qy
1141 ATGCTGCTCAGTGACAGCTTTGCTCTCTCCCATCTTAAATGGGCCCACTCTTCAAG 1200
Db
1120 ACAGAGTCTAGACTCGTGGTGGACTTCTCAATTTCTAGGGGGAGCACCACTGTCC 1179
Qy
1201 AGAGGATGAAGTCTTATAAACTGCACCTGATCACCTTTTATCTCAATGTGTGTAGTTC 1260
Db
1180 TGGCCAAATTCGCAGTCCCAACCTCAATCACTACCAACCTTCTGTCTCCAAATTTG 1239
Qy
1261 TCGTGCCCATCATTTGGCAATAGTGGCAGCTCAGCTCTGAAATGGGAAACAGAAATTCGA 1320
Db
1240 TCCCTGGCTATC-----GCTGATGTGTCTGCGCGTTTATC 1276
Qy
1321 CGGTGGCTCAGTTAATCAGATATATCTCCAAGTCCGGAAGGCCAAGGAATGCGAGTG 1380
Db
1277 ATATTTCTCTTCACTCTCTGCTATGCTCATCTTCTTGTGGTCTTCTGGACTACCAA 1336
Qy
1381 AAGATGAATCAGATTTTCAGAGAGCTGTGAGAAAGCATGAGCAACATGAAAGACAGAA 1440
Db
1337 GGTATGTGGCGGTTGTCTCTACTTCCAGAA---CATCACTACAGCACCGGACCA 1393
Qy
1441 TCCAGTATCTTTTCAGATTAATGAAGCAATCTCTAGATGCTAAGAAATTTCCAAATTTCA 1500
Db
1394 TGCAAGACCTGCACGATTCCTGCTCAAGGAACCTCTATGTTTCCCTCTT-----GTTGCT 1448
Qy
1501 GCATGAACATGATCAAGATTTAATGATGTCTTTTCCAGCTAAATCTCTTACTTCTCT 1560
Db
1449 GTACAAAACCTTCGGACGGAACCTGCACCTGTATTTCCATCCCATCATCTCTGGGCTTTCG 1508
Qy
1561 CCATCCAGGAACATGAGAATATCATAGGGGATATCTCCAAGTCTATTAGTAGTCTGAACA 1620
Db
1509 CAAGATTCCTATGGGAGTGGGCTCAGTCCGTTTCTCTCTCG----- 1549
Qy
1621 CCACAGTACTTGAATTTGCAGTTTCAGTATTTGAAACACTGAATGGCAGAGTCCAGAGAAATG 1680
Db
1550 -CTCAGTTTACTAGTGCCATTGTTTCACTGGTTC---GTAGGGCTTTTCCCCCACTGTTG 1605
Qy
1681 CATTTAAACACAGAGAGATGCGTAAATTAGAGAGCGTATATACATGCAATCAGCAG 1740
Db
1606 GCTTTCACTTATATG-GATGATGTGGTATTTGGGGGCCAAGTCTGTACAACTCTTGAGTC 1664
Qy
1741 AAAATTAAGTCTCTAGATGAAAAACAAGTATATTTTGGAAACAGAAATAAAGGGGAAATGA 1800
Db
1665 CTTTTHACCTCTATTACCAATTTTCTTGTCTTTGGGTATACATTTAAACCCCTAATAA 1724
Qy
1801 AACTGTTGAATATATCACTAATGATCTGAGGCTGAAGGATTTGGGAACATTTCTCAGACAT 1860
Db
1725 AAC-----CAAGGT 1734
Qy
1861 TGAATAATATCACTTTACTCAGGTGCCAGAAAGTGTCTGCTACTGCGGAATGGACCA 1920
Db
1735 TGGGGCTACTCCCTTAACCTTCATGGGATATGTAATTTGGATGTGGG----- 1781
Qy
1921 ACGATCTGGGCTCCAAACATGACCATCGGGGCTGTGNAACAGCAGAGGTGAATTCACAGCA 1980

1782 ----- 1781
Qy
1981 CCTACATCAAGCGTAAACAGCCACATCAAAATGAGATCAAAAGATCACCACCTGATGGGA 2040
Db
1782 -----TACTTTACCGCAAGAACATATTGTACTAAAAATCAAGCAATGTTTCG- 1829
Qy
2041 CACAAAACACATCAACAAAGAGGACCCAGCCCACTTTGGCTTCCAGCTCAATTTGGAAGT 2100
Db
1830 -----AAAACCTGCTGTAAATAGACCTATTGATTGGAAGATATGTCAGAGACTTTGTGGGT 1884
Qy
2101 TTTTCAGAGTCCACCACTGTCTTTCACGGGCCAGTCTTTCATAGACAGAAATGGGAAGAGG 2160
Db
1885 CTTTGGGCTTTGCTGCCCTTTTACAAATGTGGCTATCTTGCCTTAATGCTTTATAT 1944
Qy
2161 TCTGAAAGACATGTGGCTCTGCGGTCAAGTGTAAATGACATTGGTGTAGTACTGGAAG 2220
Db
1945 GCATGTATACAACT----- 1959
Qy
2221 CTACAGGCTGGCATCAACATCTTCACTCGCTCGCCGCAACACAGAGAGTGTAGTGTG 2280
Db
1960 -----AAGCAGGCTTTCACTTTCTGCGCAACTTACAGGCCCTTCTGTGTGA 2005
Qy
2281 ACCAAGTCTCTCTGGACTCCAGGTGAAAAAGGAGATAGAGGCCCTCTCGACAAAAATGG 2340
Db
2006 AACAAATATCTGAACCTTTTACCCCGTTGCCCGCA----- 2039
Qy
2341 TATACAGGCTTTCCAGGTCTAATAGGTACTTCCAGGTCTTAAAGGTGATCGGGGGGATCT 2400
Db
2040 ----- 2039
Qy
2401 CTGCTTTACCTGGAGTTTCAGGATTTCACAGGACCAATGGGGAGACCCGGGAAGCCAGGAC 2460
Db
2040 ----- 2039
Qy
2461 TTAATGGAACAAAAGGCCAGAGGAGAAAAAGGGAGTGGGAAGCATGCAAAAGACAATCTA 2520
Db
2040 -----ACGGTCAG 2047
Qy
2521 ATACAGTCCACTGTGTGGTGGCAGCGGCCCTCACGAAGCAGAGTGTGAGATTTTTCACG 2580
Db
2048 GTCTCTGCCAAGTGTGTGTGTCGACGCAACCCCACTGATGGGGCTTGGCTATCGGCCATA 2107
Qy
2581 AAGGCCAGTGGGTAACGCTGTGACGACCGCTGGGAACCTGCGTGGAGGACTGTGCTCT 2640
Db
2108 GCGCATGCGCGGACCTTTGTGGCTCTCTGCGG----- 2141
Qy
2641 GCAGGAGCTTGGGATACAAAGGTGTTCAAAAGTGTGCAATAAGCGAGCTTATTTTGGAAAAAG 2700
Db
2142 -----ATCCATCTGCGGAACTCTTACGAGCTTGT----- 2171
Qy
2701 GTAGGGTCCAATATGCTGAATGAAGTATTTTGTGGGAAAGAGTCAATCCATTGAAG 2760
Db
2172 ----- 2171
Qy
2761 AGTCAGAAATTAGACAGTGGGTGTGAGAGCCTGTTTCGACGACGAGAGATCTCGGGGTC 2820
Db
2172 -----TTTGTCTGCGAGCGGCTGAGGCGAA 2197
Qy
2821 ACTTTGCACCTTACATAATGATCATATTTTCAATTCACATTTTAAAACTGTTTAAAAAGTG 2880
Db
2198 ACTTATCGGC----- 2207
Qy
2881 ATTTTTCCTTTGCTTCACTAAAAATCAGCTTAATTAATTAATAAGAACTAAGAAATTTT 2940
Db
2208 ----- 2207
Qy
2941 ATCCACAGAAAAAGAAATTTTAAAAATCACTGGGTAACATATAAAATAGCTTCATATTT 3000
Db
2208 -----ACGCAACTCT 2219
Qy
3001 GCTTCAAAATACCAGAACCAATTTCAACTCTCTAGGTTTTTAAAGTGGCTCGTGGCGAAATG 3060

```
Db 2220 GTTGTCTCTCTCGGAATACACCTCCCTTTCATGGCTGCTAGGTGTGCTGCCAATGG 2279
Qy 3061 ATCCCTCAGGATATAGTAGTTTCCTTTTGTGATAGGAGGGGAAATGTAGTCTTATGC 3120
Db 2280 ATCCCTCAGGATATAGTAGTTTCCTTTTGTGATAGGAGGGGAAATGTAGTCTTATGC 2339
Qy 3121 AATACTCTTGTAGTCTTCAACATGTTAAGTATAGTATAGTATAGTATAGTATAGTATAG 3180
Db 2340 AATACACTTGTAGTCTTCAACATGTTAAGTATAGTATAGTATAGTATAGTATAGTATAG 2399
Qy 3181 GAAAGAAGCACCGTGTGATCCGATTTGGTGAAGTAAAGTGAAGTGAAGTGAAGTGAAGTGA 3240
Db 2400 GAAAGAAGCACCGTGTGATCCGATTTGGTGAAGTAAAGTGAAGTGAAGTGAAGTGAAGTGA 2459
Qy 3241 AAGGCAACAGACGGGTGTGATGATTTGGTGAAGTAAAGTGAAGTGAAGTGAAGTGAAGTGA 3300
Db 2460 AAGGCAACAGACGGGTGTGATGATTTGGTGAAGTAAAGTGAAGTGAAGTGAAGTGAAGTGA 2519
Qy 3301 -ATTGTATTTAAGTGCCTAGTCTGATACAGCAAGCCCA--TTTGACCAATTCACCAATTT 3357
Db 2520 AATTGTATTTAAGTGCCTAGTCTGATACAGCAAGCCCAATTTTGACCAATTCACCAATTT 2579
Qy 3358 GGTGTGCACCT--CCAAAGCTTCAACGCTGCGCAAGCACTCAGGGCGCAAGGGCTGCTAAA 3415
Db 2580 GGTGTGCACCTTCCAAAGCTTCAACGCTGCGCAAGCACTCAGGGCGCAAGGGCTGCTAAA 2639
Qy 3416 GGAAGCGAAACAGTGAAGAACAGTCCGCAAGAAACGGTGTGACACCCCGGATGAATGCA 3475
Db 2640 GGAAGCGAAACAGTGAAGAACAGTCCGCAAGAAACGGTGTGACACCCCGGATGAATGCA 2699
Qy 3476 GCTACTGGGCTATCTGGACAGGGGAAACGCAAGCGCAAGAAAGCAGAGTACCTTGCA 3535
Db 2700 GCTACTGGGCTATCTGGACAGGGGAAACGCAAGCGCAAGAAAGCAGAGTACCTTGCA 2759
Qy 3536 GTGGCTTTACATGGCGATAGTACTGGCGGTTTATGACAGCAGCAGCAACCGGAAT 3595
Db 2760 GTGGCTTTACATGGCGATAGTACTGGCGGTTTATGACAGCAGCAGCAACCGGAAT 2819
Qy 3596 TGCCAGCTGGGGCGCCCTCTGGTAAAGTTGGGAAGCCCTGCAAAAGTAACTGGATGGCTT 3655
Db 2820 TGCCAGCTGGGGCGCCCTCTGGTAAAGTTGGGAAGCCCTGCAAAAGTAACTGGATGGCTT 2879
Qy 3656 TCTTGCCGCCAAGATCTGATGGCGCAGGGATCAAGATCTGATCAAGACAGCAGATGAG 3715
Db 2880 TCTTGCCGCCAAGATCTGATGGCGCAGGGATCAAGATCTGATCAAGACAGCAGATGAG 2939
Qy 3716 GATCGTTTCGATGATCAACAGATGATTTGCAAGCAGGTTCTCCGCGCGCTGGGTGG 3775
Db 2940 GATCGTTTCGATGATGAAACAGATGATTTGCAAGCAGGTTCTCCGCGCGCTGGGTGG 2999
Qy 3776 AGAGGCTATTCGGCTATGACTGGGCAACAGACAAATCGGCTGCTCTGATGCCCGCGTGT 3835
Db 3000 AGAGGCTATTCGGCTATGACTGGGCAACAGACAAATCGGCTGCTCTGATGCCCGCGTGT 3059
Qy 3836 TCCGGCTGTACGCGCAGGGCGCCCGTCTTTTGTCAAGACCGACCTGTCGGGTGCC 3895
Db 3060 TCCGGCTGTACGCGCAGGGCGCCCGTCTTTTGTCAAGACCGACCTGTCGGGTGCC 3119
Qy 3896 TGAATGAATCGACGACGAGCGCGGCTATCGTGGCTGGCCACGAGCGGCGTCTCTT 3955
Db 3120 TGAATGAATCGACGACGAGCGCGGCTATCGTGGCTGGCCACGAGCGGCGTCTCTT 3179
Qy 3956 GCGCAGCTGTGCTCGACGCTGTCTACTGAAGCGGGGAGGGAATGCTGCTGATTTGGCGAAG 4015
Db 3180 GCGCAGCTGTGCTCGACGCTGTCTACTGAAGCGGGGAGGGAATGCTGCTGATTTGGCGAAG 3239
Qy 4016 TGCGGGGCGAGATCTCTGTGATCTCACCTTGTCTCTGCGGAGAAAGTATCCATCATGG 4075
Db 3240 TGCGGGGCGAGATCTCTGTGATCTCACCTTGTCTCTGCGGAGAAAGTATCCATCATGG 3299
Qy 4076 CTGATGCAATGCGGGCGGTGATACGCTTGTATCGGCTACCTGCGCATTCACACCAAG 4135
Db 3300 CTGATGCAATGCGGGCGGTGATACGCTTGTATCGGCTACCTGCGCATTCACACCAAG 3359
```

```
Qy 4136 CGAAACATCGCATCGAGCGAGCACTTCTCGATGGAAGCCGGTCTTGTGATCAGGATG 4195
Db 3360 CGAAACATCGCATCGAGCGAGCACTTCTCGATGGAAGCCGGTCTTGTGATCAGGATG 3419
Qy 4196 ATCTGGACGAAGACATCAGGGCTCGCGCAGCGCAACTGTTCCCAAGGCTCAAGGCGC 4255
Db 3420 ATCTGGACGAAGACATCAGGGCTCGCGCAGCGCAACTGTTCCCAAGGCTCAAGGCGC 3479
Qy 4256 GCATGCCCGACGCGGAGGATCTCGTCTGTGACCCATGCGCATGCCCTGTGCGCAATATCA 4315
Db 3480 GCATGCCCGACGCGGAGGATCTCGTCTGTGACCCATGCGCATGCCCTGTGCGCAATATCA 3539
Qy 4316 TGGTGGAAATGCGCGCTTTTCTCGATTCATCGATGTGCGCGGCTGGGTGTGCGGAC 4375
Db 3540 TGGTGGAAATGCGCGCTTTTCTCGATTCATCGATGTGCGCGGCTGGGTGTGCGGAC 3599
Qy 4376 GCTATCAGGACATAGCTTGGCTACCGCTGATATTGCTGAAGAGCTTGGCGCGCAATGG 4435
Db 3600 GCTATCAGGACATAGCTTGGCTACCGCTGATATTGCTGAAGAGCTTGGCGCGCAATGG 3659
Qy 4436 CTGACCGCTTCTCTGCTTTTACGCTATCGCGCTCCCGATTCGCGAGCATCGCTTCT 4495
Db 3660 CTGACCGCTTCTCTGCTTTTACGCTATCGCGCTCCCGATTCGCGAGCATCGCTTCT 3719
Qy 4496 ATCCCTTTCTGACGAGTCTTCTTGAAGCGGACTCTGGGGTTCGATA 4542
Db 3720 ATCCCTTTCTGACGAGTCTTCTTGAAGCGGACTCTGGGGTTCGATA 3766
```

RESULT 7

```
US-11-018-895-9
; Sequence 9, Application US/11018895
; Publication No. US20050100952A1
; GENERAL INFORMATION:
; APPLICANT: Bremel, Robert
; APPLICANT: Miller, Linda
; APPLICANT: Bleck, Gregory
; TITLE OF INVENTION: Host Cells Containing Multiple Integrating Vectors
; FILE REFERENCE: GALA-06416
; CURRENT APPLICATION NUMBER: US/11/018,895
; CURRENT FILING DATE: 2004-12-21
; PRIOR APPLICATION NUMBER: US/09/897,511
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 60/215,925
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 5130
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-11-018-895-9
```

```
Query Match 30.7%; Score 1588.6; DB 10; Length 5130;
Best Local Similarity 66.1%; Pred. No. 0;
Matches 3007; Conservative 0; Mismatches 754; Indels 786; Gaps 18;

Qy 1 TTTGAAAGACCCACCCCTAGGTGGCAAGCTTAGTTAAGTAAACGCCACTTTTGAAGGCAT 60
Db 1 TTTGAAAGACCCACCCCTAGGTGGCAAGCTTAGTTAAGTAAACGCCACTTTTGAAGGCAT 60
Qy 61 GGAATAATACATACTGAGATAGAAAAGTTTCAAGTCAAGTCAAGACAAAGAACAGC 120
Db 61 GGAATAATACATACTGAGATAGAAAAGTTTCAAGTCAAGTCAAGACAAAGAACAGC 120
Qy 121 TGAATACCAACAGGATATCTGTGTAAGCGGTTCTTCCCGCGGCTCAGGCGCAAGAAC 180
Db 121 TGAATACCAACAGGATATCTGTGTAAGCGGTTCTTCCCGCGGCTCAGGCGCAAGAAC 180
Qy 181 GATGAGACAGTGAAGTGAAGCGCCAAACAGGATATCTGTGTAAGCAGATTCCTGCCCCGG 240
```

Db GATGAGACAGCTGAGTGATGGCCAAAACAGGATATCTGTGTAAAGCAGTCTCTGCCCGG 240
Qy CTCGGGCGAAGAACAGATGCTCCAGATGGGTCCAGCCCTCAGAGTTTCTAGTGAA 300
Db CTCGGGCGAAGAACAGATGCTCCAGATGGGTCCAGCCCTCAGCAGTTTCTAGTGAA 300
Qy TCATCAGATGTTTCCAGGCTGCCCAAGGACTGAAATGACCCCTGTACCTTATTGAAC 360
Db TCATCAGATGTTTCCAGGCTGCCCAAGGACTGAAATGACCCCTGTACCTTATTGAAC 360
Qy TAACCAATCAGTCTCTCGCTTCTGCTTCTGCTGCGGCTTCCGCTCTCCGAGCTCAATAAA 420
Db TAACCAATCAGTCTCTCGCTTCTGCTTCTGCTGCGGCTTCCGCTCTCCGAGCTCAATAAA 420
Qy AGAGCCCAACACCTCTACTGGGCGCCAGTCTTCGATAGACTGGTCCCGGGTAC 480
Db AGAGCCCAACACCTCTACTGGGCGCCAGTCTTCGATAGACTGGTCCCGGGTAC 480
Qy CCGTATTCCTCAATAAGCCTCTTGTCTGTTTCGATCCGAATCGTGTCTCGCTGTTCTTGG 540
Db CCGTATTCCTCAATAAGCCTCTTGTCTGTTTCGATCCGAATCGTGTCTCGCTGTTCTTGG 540
Qy GGAGGGTCTCTCTGAGTGATTGACTACCCAGCGGGGCTCTTTCATTTGGGGGCTCGT 600
Db GGAGGGTCTCTCTGAGTGATTGACTACCCAGCGGGGCTCTTTCATTTGGGGGCTCGT 600
Qy CCGGATTTGAGAGCCCTGCGCAGGAGACCAACGACCCAGCGGAGGTAAAGTGGCC 660
Db CCGGATTTGAGAGCCCTGCGCAGGAGACCAACGACCCAGCGGAGGTAAAGTGGCC 660
Qy AGCACTTATCTGTGCTGCTCGATGTTGCTAGTGTCTATGTTTATGTTATGCGCCCTGG 720
Db AGCACTTATCTGTGCTGCTCGATGTTGCTAGTGTCTATGTTTATGTTATGCGCCCTGG 720
Qy TCTGTACTAGTTAGTAACTAGTCTGTATCTGGCGACCCGCTGTGGAACCTGACGAGTT 780
Db TCTGTACTAGTTAGTAACTAGTCTGTATCTGGCGACCCGCTGTGGAACCTGACGAGTT 780
Qy CTGAACACCCCGCGCAACCCCTGGAGACGTCCTGAGGACTTTGGGGCCGTTTGTGG 840
Db CTGAACACCCCGCGCAACCCCTGGAGACGTCCTGAGGACTTTGGGGCCGTTTGTGG 840
Qy CCCGACCTGAGAGGAGTGCATGTGAATCCGACCCCGTCAGATATGTTCTGCT 900
Db CCCGACCTGAGAGGAGTGCATGTGAATCCGACCCCGTCAGATATGTTCTGCT 900
Qy AGGAGACGAGAACCTTAAACAGTTCCCGCTCCGCTGGAATTTTGTCTTTCGTTTGGAA 960
Db AGGAGACGAGAACCTTAAACAGTTCCCGCTCCGCTGGAATTTTGTCTTTCGTTTGGAA 960
Qy CCGAAGCCGCGGCTTGTCTGTCGACGCAAGCTTGGGCTGCGAGTCTAGAGGA 1020
Db CCGAAGCCGCGGCTTGTCTGTCGACGCAAGCTTGGGCTGCGAGTCTAGAGGA 1020
Qy TCAATTCCGACGAGTAAATCGGTCTCGCTCTTATAGGACATATGAGTATGCCACAGT 1080
Db TCAATTCCGACGAGTAAATCGGTCTCGCTCTTATAGGACATATGAGTATGCCACAGT 1080
Qy GGGATGACTTTTCTGATCAGCAAGAGGACACTGACAGCTGTACAGAGTCTGTGAAGTTG 1140
Db GGGATGACTTTTCTGATCAGCAAGAGGAGTTTCTGTGTGACAGATCTCTCACATACC 1119
Qy ATGCTCGCTCAGTGACAGCTTTGCTCTCCCAATCTTAAATATGAGGCACTCTTCAAG 1200
Db ACAGAGTCTAGACTCGTGGTGAATCTCTCAATTTCTAGGGGAGCACCAAGTGTCC 1179
Qy AGAGATGAAGTCTTATATAAATGCACTGATGACCTTATCTCATTTGTGTTGTAGTTC 1260
Db TGGCCAAATTTGGCAGTCCCCCAACTTCAATCACTACCAACCTCTTGTCTCCAAATTG 1239
Qy TCGTGCCCATCAATTGGCATAGTGGCAGCTCAGCTCCTGAAATGGGAAACGAATTGCA 1320

Db TCCTGGCTATC-----GCTGGATGTGTCTGGCGGCTTTTATC 1276
Qy CGGTTGGCTCAGTTAATGTCAGATATATCTCCAAGTCCGGAAGGCAAGGAATGGCAGTG 1380
Db ATATTCTCTTCATCTCTGCTATGCTCATCTTCTTGTGTCTTCTGGAATCAACA 1336
Qy AAGATGAATGAGATTTTCGAGAGCTGTGATGGAAACGATGAGCAACATGAAAGCAGAA 1440
Db GGTATGTTGGCCGTTTGTCTCTACTTCCAGGA---CATCACTACAGCACGGACCA 1393
Qy TCCAGTATCTTTCAGATAATAAGCCAATCTCTAGATGCTAAGAAATTTCCAAATTTCA 1500
Db TGCAAGACCTGCACGATTTCTGCTCAAGGAACCTCTATGTTTCCCTCTT-----GTTGCT 1448
Qy GCATAACAATGATCAAGATTTAATGATGTTCTTTTCCAGCTAAATCTTACTTCTTCT 1560
Db GTACAAACCTTCGGACGGAACCTGCACCTTGTATTCCTCATCCCATCTCTGGGCTTTCG 1508
Qy CCATCCAGGAACATGAGATATCATAGGGATATCTCCAAGTCAATTAGTAGGTCTGAACA 1620
Db CAAGATTCCTATGGAGTGGGCTCAGTCCGTTTCTCTCG----- 1549
Qy CCACAGTACTTGAATTTGCAGTTTGAACACTGAAATGGCAGAGTCCAAGAGATG 1680
Db -CTCAGTTTACTAGTGCAATTTGTTCACTGTTTCT---GTAGGGCTTTTCCCACTGTTG 1605
Qy CATTTAAACAACAGAGGAGATGCGTAAATTAGAGGAGCGTATATACATGCATCAGCAG 1740
Db GCTTTCAGTTATATG--GATGATGTTGTTTGGGGCCCAAGTCTGTACAACATCTTGAGTC 1664
Qy AATTTAAGTCTCTAGATGAAAAACAAGTATATTTGGAACAGGAATAAAGGGGAATGA 1800
Db CTTTTTACCTCTATTAACCAATTTTCTTTGTTCTTGGGTATACATTTAAACCTTAATA 1724
Qy AACTGTTGAAATATACATTAATCTGAGGCTGGAAGGATTTGGGAACATTTCTCAGACAT 1860
Db AAC-----CAACGTT-----CAACGTT 1734
Qy TGAATAATATCATTTTACTCCAAGGTGCCAAGATGCTCGCTGACTGGGAAATGGACCA 1920
Db TGGGGCTACTCCCTTAACCTTCACTGGGATATGTAATTTGGATGTTGGGG----- 1781
Qy ACGATCTGGGCTCCAACATGACCATCGGGCTGTGAACAGCAGAGGTTGAATTCACAGCA 1980
Db ----- 1782
Qy CCTACATCACAGCGGTAAACAGCCCATCAAAATGAGATCAAAAGATCACCACTGCATGGGA 2040
Db -----TACTTTACCGCAAGAACATATTTGTAATAAATCAAGCAATGTTTTCG- 1829
Qy CACAAACACATCAACAGAGGAGCCAGCCACCTTTGGCTTCAACGTCATTTGGGAAGT 2100
Db -----AAAACGTGCTGTAAATAGACCTTATGTTGGAAGTATGTGAGACTTTGGGT 1884
Qy TTTTCAGAGTCCACCACTGCTTTCACGGGCCAGTCTTCATAGACAGGAATGGGAAGGAG 2160
Db CTTTTGGGCTTTGTCGCCCTTTTACCAATGTGGCTATCCTGCTTAAATGCTTTATAT 1944
Qy TCCTGAAGACCATGTGGCTGCTGCGGTCAAGTGTAAATGACATTTGGTGAATGCTGGAAAG 2220
Db GCATGTATACATCT----- 1945
Qy CTACAGGGTGGCATCAACATCTTCACTCGCTGCCACACAGAGGATGAGTGAAGT 2280
Db -----AAGCAGGCTTTCACTTTCTCGCCAACTTACAAAGCCTTTCTGTGTGA 2005
Qy ACCAAGGTCTCTCTGGACTCCAGTGAAAAAGGAGATAGAGGCCCTCTCTGCAAAAATGG 2340
Db AACHATATCTGAACCTTTTACCCCGTTGCCCGCA----- 2039
Qy TATACAGGGCTTTTCCAGGCTTAATAGGTACTCCAGGTCTTTAAAGGTGATCGGGGGATCT 2400
Db ----- 2040

QY 2401 CTGGTTTACCTGGAGTTTCGAGGATTCACGAGCAATCGGGAAGCCGGGAAGCCAGGAC 2460
Db 2040 ----- -2039
QY 2461 TTAATGACAAAAGGCCAGAAAGGAGAAAGAGGAGTGGGAAGCATGCAAGACAAATCTA 2520
Db 2040 -----ACGTGAG 2047
QY 2521 ATACAGTCCGACTGGTGGGTGGCAGCGCCCTCAGAAAGCAGAGTGGAGATTTTTCACG 2580
Db 2048 GTCTCTGCAAGTGTTCGTGACGCAACCCCACTGATGGGGCTTGGCTATCGGCCATA 2107
QY 2581 AAGGCCAGTGGGTACGGTGTGTGACGACCCCTGGGAACCTCGGTGGAGGACTGGTCTGTCT 2640
Db 2108 GCCGCATCGCGGAGCTTTGTGGCTCTCTGCCG ----- 2141
QY 2641 GCAGGAGCTTGGGATACAAAGGTGTCAAAGTGTGCATPAAGCGAGCTTATTTTGGMAAAG 2700
Db 2142 -----ATCCATCTCGGGAACCTCTAGCAGCTTGT ----- 2171
QY 2701 GTACGGGTCCAATATGCGTGAATGAAGTATTTTGTTCGGGAAGAGTCAATCCATTTGAAG 2760
Db 2172 ----- 2171
QY 2761 AGTCAGAAATTAGACAGTGGGGTGTGAGAGCCTGTTCGACGACGAAGATGCTGGGGTTC 2820
Db 2172 -----TTTGTCTCGCAGGGGGTCTGGAGCGAA 2197
QY 2821 ACTTTGCACCTACATAATGATCATATTTTCAATTCACATTTTAAAGTGTATAAAGTG 2880
Db 2198 ACTTATCGGC ----- 2207
QY 2881 ATTTTTTTTCTTGTCTTCACTAAATCAGCTTAATTAATTAATTAAGAAACATAAGAAATTT 2940
Db 2208 ----- 2207
QY 2941 ATCCACAGAAAGGAATATTTAAATACTGGATAAACAATAAATAGCTTCATATTT 3000
Db 2208 -----ACCGACAACCTCT 2219
QY 3001 GCTTCAATACCAGAACCAATTTCACTTCTAGGTTTTAAAGTGGCTCGTGGCGAATYG 3060
Db 2220 GTTGTCTCTCTCGGAATAACACTCTCTTTCATGGTGTGAGGGTGTGCGCAACTGG 2279
QY 3061 ATCCCTCTCAGGATATAGTATTTTCTGCTTTTCATAGGAGGGGAAATGTAGTCTTATGC 3120
Db 2280 ATCCCTCTCAGGATATAGTATTTTCTGCTTTTCATAGGAGGGGAAATGTAGTCTTATGC 2339
QY 3121 AATACTCTTGTAGTCTTTCGCAACATGGTAAACGATGAGTTAGCAACATGCTTTTACAAAGAGA 3180
Db 2340 AATACTCTTGTAGTCTTTCGCAACATGGTAAACGATGAGTTAGCAACATGCTTTTACAAAGAGA 2399
QY 3181 GAAAGACACGTCATGCCGATTTGGTGAAGTAAAGTGGTACGATCGTGGCTTATTAGG 3240
Db 2400 GAAAGACACGTCATGCCGATTTGGTGAAGTAAAGTGGTACGATCGTGGCTTATTAGG 2459
QY 3241 AAGGCAACAGACGGGTCTGACATGGATTGGACGAAACCACTGAAATTCGCAATTGACAGAGAT 3300
Db 2460 AAGGCAACAGACGGGTCTGACATGGATTGGACGAAACCACTGAAATTCGCAATTGACAGAGAT 2519
QY 3301 -ATTGTATTTAAGTGCCTAGTCTGATACAGCAAAAGGCCA - -TTTGACCAATTCACCAATTT 3357
Db 2520 AATTGTATTTAAGTGCCTAGTCTGATACAGCAAAAGGCCAATTTTGACCAATTCACCAATTT 2579
QY 3358 GGTGTGCACTT - CCAAGCTTTCAGCTGCGCGCAAGCACTCAGGGCGCAAGGGCTGCTAAA 3415
Db 2580 GGTGTGCACTTTCAAAAGCTTTCAGCTGCGCGCAAGCACTCAGGGCGCAAGGGCTGCTAAA 2639
QY 3416 GGAAGCGGAACACCTAGAAAGCCAGTCCGAGAAACCGTGTGACCCCGGATGAATGTCA 3475
Db 2640 GGAAGCGGAACACCTAGAAAGCCAGTCCGAGAAACCGTGTGACCCCGGATGAATGTCA 2699

QY 3476 GCTACTGGGCTATCTGACAAAGGGAACCGCAAGCCAAAGAGAAAGAGAGTAGCTTGCA 3535
Db 2700 GCTACTGGGCTATCTGCAAGGGGAAACCGCAAGCCAAAGAGAAAGAGAGTAGCTTGCA 2759
QY 3536 GTGGGCTTTACATGCGGATAGCTAGACTGGGCGGTTTTATGACAGCAAGCAACCGGAAT 3595
Db 2760 GTGGGCTTTACATGCGGATAGCTAGACTGGGCGGTTTTATGACAGCAAGCAACCGGAAT 2819
QY 3596 TGCCAGCTGGGGGCCCTCTGTTAAGGTTGGGAAGCCCTGCAAAAGTAAACTGGATGGCTT 3655
Db 2820 TGCCAGCTGGGGGCCCTCTGTTAAGGTTGGGAAGCCCTGCAAAAGTAAACTGGATGGCTT 2879
QY 3656 TCTTGGCCCAAGAGATCTGATGGCGCAGGGGATCAAGATCTGATCAAGAGACAGATGAG 3715
Db 2880 TCTTGGCCCAAGAGATCTGATGGCGCAGGGGATCAAGATCTGATCAAGAGACAGATGAG 2939
QY 3716 GATCGTTTTCGATGATTAACAAGATGATTTGACGACGAGGTTCTCCGGCCGCTTGGGTGG 3775
Db 2940 GATCGTTTTCGATGATTAACAAGATGATTTGACGACGAGGTTCTCCGGCCGCTTGGGTGG 2999
QY 3776 AGAGGCTTATTCGGCTATGACTGGGCACAAACAGACAATCGGCTGCTCTGATGCCCGCGTGT 3835
Db 3000 AGAGGCTTATTCGGCTATGACTGGGCACAAACAGACAATCGGCTGCTCTGATGCCCGCGTGT 3059
QY 3836 TCCGGCTGTACGCGCAGGGGGCCCGGTTCTTTTGTCAAGACCGACTCTGTCCGGTGGCC 3895
Db 3060 TCCGGCTGTACGCGCAGGGGGCCCGGTTCTTTTGTCAAGACCGACTCTGTCCGGTGGCC 3119
QY 3896 TGAATGAACCTGCAGACGAGCGCGGCTATCGTGGCTGGCCACGACGCGGCTTCCCTT 3955
Db 3120 TGAATGAACCTGCAGACGAGCGCGGCTATCGTGGCTGGCCACGACGCGGCTTCCCTT 3179
QY 3956 GCGCAGCTGTCTCGAAGCTTGTCACTGAAGCGGGAAGGAGCTGGCTGTCTATTGGGCGAAG 4015
Db 3180 GCGCAGCTGTCTCGAAGCTTGTCACTGAAGCGGGAAGGAGCTGGCTGTCTATTGGGCGAAG 3239
QY 4016 TGCCGGGCGAGGATCTCTCTGTCACTCACTTGTCTGCTGCGGAGAAAGTATCCATCATGG 4075
Db 3240 TGCCGGGCGAGGATCTCTCTGTCACTCACTTGTCTGCTGCGGAGAAAGTATCCATCATGG 3299
QY 4076 CTGATGCAATGCGCGGCTGCATACGCTTGTATCGGCTACCTGCCCATTCGACCAACCAAG 4135
Db 3300 CTGATGCAATGCGCGGCTGCATACGCTTGTATCGGCTACCTGCCCATTCGACCAACCAAG 3359
QY 4136 CGAAACATCGCATCGACGAGCAGTACTCGGATGGAAGCGGCTCTTGTGATCAGGATG 4195
Db 3360 CGAAACATCGCATCGACGAGCAGTACTCGGATGGAAGCGGCTCTTGTGATCAGGATG 3419
QY 4196 ATCTGGAACGAAGCATACAGGGGCTCGCGCAGCGCAACTGTTGCGCAGGCTCAAGGCGC 4255
Db 3420 ATCTGGAACGAAGCATACAGGGGCTCGCGCAGCGCAACTGTTGCGCAGGCTCAAGGCGC 3479
QY 4256 GCATGCCCGACGAGGATCTCGTGTGACCCATGGCGATGCTGCTGCCGAATATCA 4315
Db 3480 GCATGCCCGACGAGGATCTCGTGTGACCCATGGCGATGCTGCTGCCGAATATCA 3539
QY 4316 TGGTGGAAAAATGGCCGCTTTTCTGGATTCACTCACTGTGCGCGGCTGGGTGTGCGGACCC 4375
Db 3540 TGGTGGAAAAATGGCCGCTTTTCTGGATTCACTCACTGTGCGCGGCTGGGTGTGCGGACCC 3599
QY 4376 GCTATCAGGACATAGCGGCTTGGCTACCCGTTGATATTGCTGAAGAGCTTTGGCGGCGAATGGG 4435
Db 3600 GCTATCAGGACATAGCGGCTTGGCTACCCGTTGATATTGCTGAAGAGCTTTGGCGGCGAATGGG 3659
QY 4436 CTGACCGCTTCTCGTCTTTACGGTATCCGCGCTCCCGATTTCGACGCGCATCGCCTTCT 4495
Db 3660 CTGACCGCTTCTCGTCTTTACGGTATCCGCGCTCCCGATTTCGACGCGCATCGCCTTCT 3719
QY 4496 ATCGCTTCTTGAAGGATCTTCTGAGCGGAGCTCTGGGGTTCGATA 4542
Db 3720 ATCGCTTCTTGAAGGATCTTCTGAGCGGAGCTCTGGGGTTCGATA 3766

RESULT 8
US-11-036-557-9
; Sequence 9, Application US/11036557
; Publication No. US20050221429A1
; GENERAL INFORMATION:
; APPLICANT: Black, Gregory T.
; TITLE OF INVENTION: Host Cells Containing Multiple Integrating Vectors Comprising an
; FILE REFERENCE: GALA-09479
; CURRENT APPLICATION NUMBER: US/11/036,557
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 9
; LENGTH: 5130
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-11-036-557-9

Query Match 30.7%; Score 1588.6; DB 10; Length 5130;
Best Local Similarity 66.1%; Pred. No. 0;
Matches 3007; Conservative 0; Mismatches 754; Indels 786; Gaps 18;

| | | | |
|----|-----|--|-----|
| Qy | 1 | TTTGAAGACCCACCGTAGGTGGCAAGCTAGCTTAAGTAACGCCACTTTTGCAGGCAT | 60 |
| Db | 1 | TTTGAAGACCCACCGTAGGTGGCAAGCTAGCTTAAGTAACGCCACTTTTGCAGGCAT | 60 |
| Qy | 61 | GGAAAAATACATAACTGAGATAGAAAAGTTTTCAGATCAAGGTTCAGGAAACAGC | 120 |
| Db | 61 | GGAAAAATACATAACTGAGATAGAAAAGTTTTCAGATCAAGGTTCAGGAAACAGC | 120 |
| Qy | 121 | TGAATACCAACAGGATATCTGTGTAGCGGTTCTTCCCGGCTCAGGCGCAAGAAC | 180 |
| Db | 121 | TGAATACCAACAGGATATCTGTGTAGCGGTTCTTCCCGGCTCAGGCGCAAGAAC | 180 |
| Qy | 181 | GATGAGACAGCTGAGTATGGCCAAACAGGATATCTGTGTAGCGGTTCTTCCCGG | 240 |
| Db | 181 | GATGAGACAGCTGAGTATGGCCAAACAGGATATCTGTGTAGCGGTTCTTCCCGG | 240 |
| Qy | 241 | CTCGGGGCAAGAACAGATGGTCCCGCAGATCGGTCCAGCCCTCAGCAGTTTCTAGTGA | 300 |
| Db | 241 | CTCGGGGCAAGAACAGATGGTCCCGCAGATCGGTCCAGCCCTCAGCAGTTTCTAGTGA | 300 |
| Qy | 301 | TCATCAGATGTTTCCAGGGTCCCGCAGGACCTGAAATGACCTGTACCTTATTTGAAC | 360 |
| Db | 301 | TCATCAGATGTTTCCAGGGTCCCGCAGGACCTGAAATGACCTGTACCTTATTTGAAC | 360 |
| Qy | 361 | TAAACCAATCAGTTCGCTTCTGCTTCTGTCGCGGCTTCCGCTCCTCGAGCTCAATAA | 420 |
| Db | 361 | TAAACCAATCAGTTCGCTTCTGCTTCTGTCGCGGCTTCCGCTCCTCGAGCTCAATAA | 420 |
| Qy | 421 | AGAGCCCAACACCCCTCACTCGGGCGCCAGTCTTCCGATAGACTGCGTCCCGGGTAC | 480 |
| Db | 421 | AGAGCCCAACACCCCTCACTCGGGCGCCAGTCTTCCGATAGACTGCGTCCCGGGTAC | 480 |
| Qy | 481 | CCGTATTTCCCAATAAAGCTCTTGTGTTTGCATCCGATCCGATCGTGTCTCGTCTT | 540 |
| Db | 481 | CCGTATTTCCCAATAAAGCTCTTGTGTTTGCATCCGATCCGATCGTGTCTCGTCTT | 540 |
| Qy | 541 | GGAGGCTCTCTGTAGTATGATCTCCACGACGGGGTCTTTTCAATTTGGGGGCTCGT | 600 |
| Db | 541 | GGAGGCTCTCTGTAGTATGATCTCCACGACGGGGTCTTTTCAATTTGGGGGCTCGT | 600 |
| Qy | 601 | CCGGATTTGGAGACCCCTGCGCCAGGACACCGACCCACCGAGGAGTAAAGCTGGCC | 660 |
| Db | 601 | CCGGATTTGGAGACCCCTGCGCCAGGACACCGACCCACCGAGGAGTAAAGCTGGCC | 660 |
| Qy | 661 | AGCAACTTATCTGTGTCTGTCCGATGTCATGTTGATGTTATGCGGCTCGG | 720 |
| Db | 661 | AGCAACTTATCTGTGTCTGTCCGATGTCATGTTGATGTTATGCGGCTCGG | 720 |

| | | | |
|----|------|---|------|
| Qy | 721 | TCGTACTAGTTAGCTAACTAGCTCTGTATCTGGCGACCCGCTGGTGGAACTGACGAGTT | 780 |
| Db | 721 | TCGTACTAGTTAGCTAACTAGCTCTGTATCTGGCGACCCGCTGGTGGAACTGACGAGTT | 780 |
| Qy | 781 | CTGAACACCCCGCGCAACCTCTGGAGACGCTCCAGGACCTTTGGGGCCGCTTTTGTGG | 840 |
| Db | 781 | CTGAACACCCCGCGCAACCTCTGGAGACGCTCCAGGACCTTTGGGGCCGCTTTTGTGG | 840 |
| Qy | 841 | CCGACCTGAGGAGGAGTGGATGCGAATCCGACCCCGCTCAGGATATGCTTCTGGT | 900 |
| Db | 841 | CCGACCTGAGGAGGAGTGGATGCGAATCCGACCCCGCTCAGGATATGCTTCTGGT | 900 |
| Qy | 901 | AGGAGACGAGAACCTAAACAGATTCCCGCTCTGAAATTTTGTCTTTCGGTTTGAA | 960 |
| Db | 901 | AGGAGACGAGAACCTAAACAGATTCCCGCTCTGAAATTTTGTCTTTCGGTTTGAA | 960 |
| Qy | 961 | CCGAAGCCGCGCTTTGTCTGTGACCAAGCTTGGGCTGACAGGTCGAGGACTGGGA | 1020 |
| Db | 961 | CCGAAGCCGCGCTTTGTCTGTGACCAAGCTTGGGCTGACAGGTCGAGGACTGGGA | 1020 |
| Qy | 1021 | TCAAATCCGACAGGTAATCCGCTGCTCTTAGGACATATGAAGTATGACACAGT | 1080 |
| Db | 1021 | CC-----CTGCACCGAACATGGAGAACACACATCAGGATTCCT | 1059 |
| Qy | 1081 | GGGATGACTTTCTGATCAGCAAGAGGACACTGACAGCTGTACAGAGTCTGTGAAGTTCG | 1140 |
| Db | 1060 | AGGACCCCTGCTGTGTTACAGCGGGGTTTCTTGTGTGACAGAACTCTCAACAATCC | 1119 |
| Qy | 1141 | ATGCTCGCTCAGTGACAGCTTTGCTTCTCCCATCTTAAATGGGCCCAACTCTTCAAG | 1200 |
| Db | 1120 | ACAGAGCTAGAGCTCGTGTGTGACTTCTCTCAATTTTCTAGGGGGAGCACCACTGTCC | 1179 |
| Qy | 1201 | AGAGGATGAGTCTTATAAATGCACTGATCACTTATCTCTATCTGTTGTGTTAGTTC | 1260 |
| Db | 1180 | TGGCAAAATTCGAGTCCCAACCTCCAATCACTCAACACCTCTTGTCTCCAAATTTG | 1239 |
| Qy | 1261 | TCGTGCCCATCAITGGCATAGTGGCAGCTCAGCTCTGAAATGGGAAACGAAGAAATGCA | 1320 |
| Db | 1240 | TCCTGGTATC-----GCTGGATGTGCTCGCGGCTTTTATC | 1276 |
| Qy | 1321 | CGGTGGCTCAGTTAATGACAGATATATCTCAAAGTCCGGAAGGCAAGAAATGGCAGTG | 1380 |
| Db | 1277 | ATATTCCTCTTCATCTCTGCTGCTATGCTCATCTCTTGTGTTGGTCTTCTTGGACTACCA | 1336 |
| Qy | 1381 | AGATGAAATGAGATTTCCGAGAGCTGTGATGGAACCATGAGCAACATCGAAAGAGAA | 1440 |
| Db | 1337 | GGTATGTTGCCGCTTGTCTCTTACCTTCCAGGAA---CATCAACTACAGACACGGACCA | 1393 |
| Qy | 1441 | TCCAGTATCTTTCCAGATAATGAAGCCAACTCTCTAGATGCTAAGAAATTCGAAATTTCA | 1500 |
| Db | 1394 | TGCAGACCTGACAGATCTCTGCTCAAGGAACCTCTATGTTTCCCTCTT-----GTTGCT | 1448 |
| Qy | 1501 | GCATAACAACCTGATCAAGATTTAAATGATGTTCTTTTCCAGCTAAATTCCTTACTTCCCT | 1560 |
| Db | 1449 | GTACAAAACCTTCGAGCGGAACTGCACTTGTATTCCTCATCCCATCATCTCTGGGCTTCG | 1508 |
| Qy | 1561 | CCATCCAGGACATGAGAAATATCATAGGGGATATCTCCAAAGTCATTAGTAGTCTGAACA | 1620 |
| Db | 1509 | CAAGATTCCTATGGGAGTGGGCTCAGTCCGCTTCTCTCTG----- | 1549 |
| Qy | 1621 | CCAGCTACTTGTATTTGACGTTTCAAGTATGAACACATGGAATGGCAGAGTCCAAGAGAAATG | 1680 |
| Db | 1550 | -CTCAGTTTACTAGTGCCATTTGTTTCAAGTGGTTC---GTAGGGCTTTTCCCTCTTTG | 1605 |
| Qy | 1681 | CATTTAAACAACAGAGGAGATGCGTAAATTAGAGGAGCGTATATACAAATGCATCAGCAG | 1740 |
| Db | 1606 | GCITTCAGTTATATG-GATGATGTGTATTTGGGGCCCAAGTCTGTACAACTCTTGTAGTC | 1664 |
| Qy | 1741 | AAATTAAGTCTCTAGATGAAGAAACAGTATATTTTGGAAACAGGAAATTAAGGGGAAATGA | 1800 |
| Db | 1665 | CTTTTACCTCTATTAACCAATTTTCTTGTCTTGTGGGTATATCAATTTAAACCTTAATA | 1724 |
| Qy | 1801 | AACTGTTGAATATATCACTAATGATCTGAGGCTGAAGGATTTGGGAAACATTTCTCAGACAT | 1860 |

Db 1725 AAC ||||| 1734
Qy 1861 TGAATAATATCACTTTACTCAAGTGCAGAAAGTCTCCTGACTGGGAAATGGACCA 1920
Db 1735 TGGGCTACTCCCTTAACCTTATGGATATGTAATGGATGTTGGG 1781
Qy 1921 ACGATCTGGGCTCCAAATGACATGACATCGGGGCTGTGAACAGCAGAGGTGAATTCAGAGCA 1980
Db 1782 1781
Qy 1981 CCTACATCACAGCGGTAAACAGCCATCAATGAGATCAAGAGTCAACACTGCATGGGA 2040
Db 1782 TACTTTACCCGCAAGAACATATGTACTTAAATCAAGCAATGTTTCG- 1829
Qy 2041 CACAAAACACCATCAACAAAGAGGACCCAGGCCACTTTGGCTTCAACCGTCAATTTGGAAGT 2100
Db 1830 AAAAGCTCTGTAATAGACCTATTGATGGAAAGTATCTCAGAGACTTGTGGGT 1884
Qy 2101 TTTGAGTCCACCACTGTCTTTCAGGGGCCAGTCTTCATAGACAGGAATGGGAAGG 2160
Db 1885 CTTTGGGCTTTGTGCCCCCTTTTACACAATGTGGCTATCTGCTTAATGCTTTATAT 1944
Qy 2161 TCCTGAGACCATGTGGCTGCTGGGTCAAGTGTAAATGACATTTGGTGTGACTGGAAG 2220
Db 1945 GCATGTATACAACT- 1959
Qy 2221 CTACAGGGTGGGATCAACATCTTCACTCGCTCGCCACACAGAGAGTGAAGTGT 2280
Db 1960 AAGAGGCTTTCACTTTCTCGCCAACTTACAGGCCCTTTCTGTGTA 2005
Qy 2281 ACCAAGTCTCTCGACTCCAGGTGAAAAGGAGATAGAGGCCCTCTGGAACAAATGG 2340
Db 2006 AACAAATCTGAACCTTTACCCCGTTGCCCGCA- 2039
Qy 2341 TATACAGGCTTTCCAGTCTAATAGTACTCCAGGTCTTAAAGTGTATCGGGGGATCT 2400
Db 2040 2039
Qy 2401 CTGGTTTACCTGGAGTTCGAGGATTCACGAGCAATCCGGAAGACCGGAGCCAGGAC 2460
Db 2040 2039
Qy 2461 TTAATGACAAAAAGCCAGAAAGGAGGAGAAAGGAGTGAAGCATGCAGAACAACTA 2520
Db 2040 ACGGTGAG 2047
Qy 2521 ATACAGTCCGACTGGTGGGTGGCAGCGCCCTCTCAGAGGAGAGTGGAGATTTTTCAG 2580
Db 2048 GTCTCTGCAAGTGTGTCTGACGCAACCCCACTGGATGGGCTTGGCTATCGGCCATA 2107
Qy 2581 AAGCCAGTGGGTACGGTGTGACGACCCGCTGGAACTCGGTGGAGGACTGGTCTCT 2640
Db 2108 GCGCATGCGCGGACCTTTGTGGCTCTCTGCGG- 2141
Qy 2641 GCAGAGCTTGGGATACAAAGGTGTTCAAAGTGTGATCAAGCGAGCTTTATTTGGAAG 2700
Db 2142 ATCCATACTGGGAACTCTAGCAGCTGT- 2171
Qy 2701 GTACGGGTCCAAATATGGCTGAATGAATTTTGTTCGGGAAGAGTCAATTCATTAAG 2760
Db 2172 2171
Qy 2761 AGTCAGAAATAGACAGTGGGTGTGAGAGCTGTTCGACGACGAGATGCTGGGGTC 2820
Db 2172 TTTGCTCGAGCGGCTCTGGAGCGAA 2197
Qy 2821 ACTTTGACCTACATAATATGATCATATTTTCAATCAATTTTTTAACTGTTATAAGTG 2880
Db 2198 ACTTATCGG- 2207
Qy 2881 ATTTTTTCTTCTGCTTCACTAAATACAGTTAATTAATTAATTAAGAACTAAGAAATTT 2940

Db 2208 2207
Qy 2941 ATCCACAGAAAGGAATATTTAAAAATCACTGGGATAAACAATATAAAATAGCTTCATATTT 3000
Db 2208 ACCGACAACTCT 2219
Qy 3001 GCTTCAATATACCAGAACCAATTTCAAATCTCTAGTGTTTTAAAGTGGCTGCGTGGCGAATTTG 3060
Db 2220 GTTGTCTCTCTCGGAATAACACCTCTCTTCCATGGCTGTAGGTGTGCTGCCAACATGG 2279
Qy 3061 ATCCCTCAGGATATAGTGTTCGCTTTTCATAGGAGGGGAAATGTAGTCTTATGC 3120
Db 2280 ATCCCTCAGGATATAGTGTTCGCTTTTCATAGGAGGGGAAATGTAGTCTTATGC 2339
Qy 3121 AATACTCTTGTAGTCTTCAACATGGTAAACGATGAGTTAGCAACATGCTTTACAAGAGAGA 3180
Db 2340 AATACTCTTGTAGTCTTCAACATGGTAAACGATGAGTTAGCAACATGCTTTACAAGAGAGA 2399
Qy 3181 GAAAAGCACCGTGCATGCGGATTTGGTGAAGTAAAGTGGTACGATCGTGTCTTATAGG 3240
Db 2400 GAAAAGCACCGTGCATGCGGATTTGGTGAAGTAAAGTGGTACGATCGTGTCTTATAGG 2459
Qy 3241 AAGSCAACAGACGGTCTGACATGATTTGGACGAACCACTGAATTTCCGATTTGAGAGAT 3300
Db 2460 AAGSCAACAGACGGTCTGACATGATTTGGACGAACCACTGAATTTCCGATTTGAGAGAT 2519
Qy 3301 -ATTGTATTTAAGTGCCTAGCTCGATACAGCAACGCCA--TTTGACCATTTCAACACATTT 3357
Db 2520 AATTGTATTTAAGTGCCTAGCTCGATACAGCAACGCCAATTTTGAACATTTCAACACATTT 2579
Qy 3358 GGTGTGCACCT--CCAAAGCTTTCAGCTGCGCGAAGCACTAGGGCGGAGGGCTGTAAA 3415
Db 2580 GGTGTGCACCTTCCAAAGCTTTCAGCTGCGCGAAGCACTAGGGCGGAGGGCTGTAAA 2639
Qy 3416 GGAAGCGGAACACGTAGAAAGCCAGTCCGAGAAACCGTGTGACCCCGGATGAATGTCA 3475
Db 2640 GGAAGCGGAACACGTAGAAAGCCAGTCCGAGAAACCGTGTGACCCCGGATGAATGTCA 2699
Qy 3476 GCTACTGGGCTATCTGGACAAAGGAAACGCAAGCAAGAGAGAAAGCAGGTAGCTTGCA 3535
Db 2700 GCTACTGGGCTATCTGGACAAAGGAAACGCAAGCGCAAGAGAGAGAGCAGGTAGCTTGCA 2759
Qy 3536 GTGGGCTTACATGCGGATAGCTAGACTGGGCGGTTTTATGACAGAGCAAGCAACCGGAAT 3595
Db 2760 GTGGGCTTACATGCGGATAGCTAGACTGGGCGGTTTTATGACAGAGCAAGCAACCGGAAT 2819
Qy 3596 TGCCAGCTGGGGCGCCCTCTGTTAAGTTGGGAAGCCCTGCAAAAGTAACTGGATGGCTT 3655
Db 2820 TGCCAGCTGGGGCGCCCTCTGTTAAGTTGGGAAGCCCTGCAAAAGTAACTGGATGGCTT 2879
Qy 3656 TCTTCCGCAAGGATCTGATGGCGCAGGGGATCAAGATCTGATCAAGAGACAGGATGAG 3715
Db 2880 TCTTCCGCAAGGATCTGATGGCGCAGGGGATCAAGATCTGATCAAGAGACAGGATGAG 2939
Qy 3716 GATCGTTTCCGATGATTAACAGATGGAATGCAAGCGAGGTTCTCGGCGCTGTGGGTGG 3775
Db 2940 GATCGTTTCCGATGATTAACAGATGGAATGCAAGCGAGGTTCTCGGCGCTGTGGGTGG 2999
Qy 3776 AGAGGCTATTCCGCTATGATCGGCAACAGCAATCGGCTGCTCTGATGCGCGCGTGT 3835
Db 3000 AGAGGCTATTCCGCTATGATCGGCAACAGCAATCGGCTGCTCTGATGCGCGCGTGT 3059
Qy 3836 TCCGGCTGTAGCGCAGGGCGCCCGGTTCTTTTGTCAAGACCGACTGTCCCGTGGCC 3895
Db 3060 TCCGGCTGTAGCGCAGGGCGCCCGGTTCTTTTGTCAAGACCGACTGTCCCGTGGCC 3119
Qy 3896 TGAATGAATCTCAGAGACGAGCGCGGCTATCGTGGCTGGCCACGACGGCGCTTCTT 3955
Db 3120 TGAATGAATCTCAGAGACGAGCGCGGCTATCGTGGCTGGCCACGACGGCGCTTCTT 3179
Qy 3956 GCGCAGCTGTCTCGAGCTGTCTAAGCGGGAAGGAGTGTGCTGCTATTTGGCGGAG 4015
Db 3180 GCGCAGCTGTCTCGAGCTGTCTAAGCGGGAAGGAGTGTGCTGCTATTTGGCGGAG 3239


```
QY 4532 GGGGTTTCGATATAAATAAGATTTTATTAGTCTCCAGAAAAGGGGGGAATGAAGACC 4591
DB 3617 GGGGTTTCGATATAAATAAGATTTTATTAGTCTCCAGAAAAGGGGGGAATGAAGACC 3676
QY 4592 CCACCTGTAGGTTTGGCAAGCTAGCTTAAGTAACGCCATTTTGAAGGATGGAATAATA 4651
DB 3677 CCACCTGTAGGTTTGGCAAGCTAGCTTAAGTAACGCCATTTTGAAGGATGGAATAATA 3736
QY 4652 CATACTAGAAATAGAGAAATTCAGATCAAGGTCAAGGACAGATGGAACAGCTGAATATG 4711
DB 3737 CATACTAGAAATAGAGAAATTCAGATCAAGGTCAAGGACAGATGGAACAGCTGAATATG 3796
QY 4712 GGCCTAAACAGGATATCTGTGTAAAGCAGTTCTTCCGCCGCTCAGGGGCAAGAACAGATG 4771
DB 3797 GGCCTAAACAGGATATCTGTGTAAAGCAGTTCTTCCGCCGCTCAGGGGCAAGAACAGATG 3856
QY 4772 GAACAGCTGAATATGGGCAACAGGATATCTGTGTAAAGCAGTTCTTCCGCCGCTCAG 4831
DB 3857 GAACAGCTGAATATGGGCAACAGGATATCTGTGTAAAGCAGTTCTTCCGCCGCTCAG 3916
QY 4832 GGCCTAAACAGGATATGGGCAACAGGATATCTGTGTAAAGCAGTTCTTCCGCCGCTCAG 4891
DB 3917 GGCCTAAACAGGATATGGGCAACAGGATATCTGTGTAAAGCAGTTCTTCCGCCGCTCAG 3976
QY 4892 AGATGTTTCCAGGGTCCCAAGGACCTGAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 4951
DB 3977 AGATGTTTCCAGGGTCCCAAGGACCTGAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 4036
QY 4952 ATCAGTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCG 5011
DB 4037 ATCAGTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCG 4096
QY 5012 CACAACCCCTCACTCGGGGCGCCAGTCTCCGATTGACTGAGTCCCGGGGTACCCGGTGT 5071
DB 4097 CACAACCCCTCACTCGGGGCGCCAGTCTCCGATTGACTGAGTCCCGGGGTACCCGGTGT 4156
QY 5072 ATCCAAATAACCCCTTCGAGTTGATCCGATTCGATTCGATTCGATTCGATTCGATTCGAT 5131
DB 4157 ATCCAAATAACCCCTTCGAGTTGATCCGATTCGATTCGATTCGATTCGATTCGATTCGAT 4216
QY 5132 CTCTCTCAGTGATGACTACCCGTCAGGGGGGTCTTCATTTGG 5177
DB 4217 CTCTCTCAGTGATGACTACCCGTCAGGGGGGTCTTCATTTGG 4262

RESULT 10
US-10-164-965A-1
; Sequence 1, Application US/10164965A
; Publication No. US20040001808A1
; GENERAL INFORMATION:
; APPLICANT: Kenneth, Haglid
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE USE OF BCL-2 TRANSFECTED NEURONS
; FILE REFERENCE: 56941-2
; CURRENT APPLICATION NUMBER: US/10/164,965A
; CURRENT FILING DATE: 2002-10-22
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 6046
; TYPE: DNA
; ORGANISM: Human Bcl-2 transfected neurons
US-10-164-965A-1

Query Match 28.6%; Score 1481; DB 6; Length 6046;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3697 GATCAAGAGACAGGATGAGGATCGTTTCGATGATTGAACAGATGATTCACGACGAGGT 3756
DB 2208 GATCAAGAGACAGGATGAGGATCGTTTCGATGATTGAACAGATGATTCACGACGAGGT 2267
QY 3757 TCTCCGGCCGCTTGGGTGGAGAGGCTATTTCGGCTATGACTCGGGCACACAGACAAATCGGC 3816
```

```
DB 2268 TCTCCGGCCGCTTGGGTGGAGAGGCTATTTCGGCTATGACTGGGCAACACAGACAAATCGGC 2327
QY 3817 TGCTCTGATCGCCGCTGTTCCGGCTGTGACGAGGGGCGCCCGGTTCTTTTGTGTCAAG 3876
DB 2328 TGCTCTGATCGCCGCTGTTCCGGCTGTGACGAGGGGCGCCCGGTTCTTTTGTGTCAAG 2387
QY 3877 ACCGACCTGTCCGGTGCCTGAAATGAATGAACTGAGAGAGGAGGAGGCGGCTATCTGTGGCTG 3936
DB 2388 ACCGACCTGTCCGGTGCCTGAAATGAATGAACTGAGAGAGGAGGCGGCTATCTGTGGCTG 2447
QY 3937 GCCACGAGGGGCTTCTTCCGCGAGCTGTGCTCGAGCTGTGTCACCTGAAAGCGGAGAGGAC 3996
DB 2448 GCCACGAGGGGCTTCTTCCGCGAGCTGTGCTCGAGCTGTGTCACCTGAAAGCGGAGAGGAC 2507
QY 3997 TGGCTGCTATTGGGCGGAAGTCCCGGGGAGGATCTCTCTGTCTCATCTCTCACTTGTCTCTGCC 4056
DB 2508 TGGCTGCTATTGGGCGGAAGTCCCGGGGAGGATCTCTCTGTCTCATCTCTCACTTGTCTCTGCC 2567
QY 4057 GAGAAAGTATCCATCATGCTGATGCAATGCGGGGCTGCAATAGCGTTGATTCGGGCTAC 4116
DB 2568 GAGAAAGTATCCATCATGCTGATGCAATGCGGGGCTGCAATAGCGTTGATTCGGGCTAC 2627
QY 4117 TGCCCATTTCCGACCAACGAAACATCGCATTCGAGCGAGCAGCTACTCTCGATGGAAGCC 4176
DB 2628 TGCCCATTTCCGACCAACGAAACATCGCATTCGAGCGAGCAGCTACTCTCGATGGAAGCC 2687
QY 4177 GGTCTGTGCGATCAGGATGATCTCGACGAGGATCAGGGGCTCGGCCAGCCGAACTG 4236
DB 2688 GGTCTGTGCGATCAGGATGATCTCGACGAGGATCAGGGGCTCGGCCAGCCGAACTG 2747
QY 4237 TTCCGCCAGGCTCAAGGGCGCATGCCGAGCGGAGGATCTCTGTGTCGACCCATGCGCAT 4296
DB 2748 TTCCGCCAGGCTCAAGGGCGCATGCCGAGCGGAGGATCTCTGTGTCGACCCATGCGCAT 2807
QY 4297 GCCTGCTTCCGGAATATCATGCTGGAATAATGCGCGCTTTCTTGATTCATCGACTGTGGC 4356
DB 2808 GCCTGCTTCCGGAATATCATGCTGGAATAATGCGCGCTTTCTTGATTCATCGACTGTGGC 2867
QY 4357 CGGCTGCTGTGGCGGACCGCTATCAGGACATAGCTGTGGCTACCCGCTGATTTGCTGAA 4416
DB 2868 CGGCTGCTGTGGCGGACCGCTATCAGGACATAGCTGTGGCTACCCGCTGATTTGCTGAA 2927
QY 4417 GAGCTTGGCGCGAATGGGCTGACCGCTTCTCTGTGCTTTACGGTATCGCCGCTCCCGAT 4476
DB 2928 GAGCTTGGCGCGAATGGGCTGACCGCTTCTCTGTGCTTTACGGTATCGCCGCTCCCGAT 2987
QY 4477 TCGCAGGCGCATCGCTTCTATCGCTTCTTGACAGGTTCTTCTGAGCGGAGTCTGGGGT 4536
DB 2988 TCGCAGGCGCATCGCTTCTATCGCTTCTTGACAGGTTCTTCTGAGCGGAGTCTGGGGT 3047
QY 4537 TCGATAAAATAAAAGATTTTATTAGTCTCCAGAAAAGGGGGGAATGAAGAGCCCAACC 4596
DB 3048 TCGATAAAATAAAAGATTTTATTAGTCTCCAGAAAAGGGGGGAATGAAGAGAGCCCAACC 3107
QY 4597 TGTAAGTTTGGCAAGCTAGCTTAAGTAACGCCATTTTTCAGAGGATGGAATAATACATAA 4656
DB 3108 TGTAAGTTTGGCAAGCTAGCTTAAGTAACGCCATTTTTCAGAGGATGGAATAATACATAA 3167
QY 4657 CTGAGAAATAGAGAAATTCAGATCAAGGTCAAGGAAACAGATGGAACAGCTGAATATGGGCCA 4716
DB 3168 CTGAGAAATAGAGAAATTCAGATCAAGGTCAAGGAAACAGATGGAACAGCTGAATATGGGCCA 3227
QY 4717 AACAGGATATCTGTGTAAGCAGTTCTTCCGCCGCTCAGGGCAACAGACAGATGGAACA 4776
DB 3228 AACAGGATATCTGTGTAAGCAGTTCTTCCGCCGCTCAGGGCAACAGACAGATGGAACA 3287
QY 4777 GCTGAATATGGGCCAACAAGGATATCTGTGTAAGCAGTTCTTCCGCCGCTCAGGGCA 4836
DB 3288 GCTGAATATGGGCCAACAAGGATATCTGTGTAAGCAGTTCTTCCGCCGCTCAGGGCA 3347
QY 4837 AGAAACAGATGCTCCCAAGATGCGGTCAGGCTCAGGCTTCTAGAGAACCATCAGATG 4896
```


Db 3348 AGAACAGATGGTCCCGCAGATGGGTCCAGCCCTCAGCAGATTCTAGAGAACCATCAGATG 3407
Qy 4897 TTTCCAGGGTCCCGCAGGACTGAATGACCCCTGTGCTTATTGTAACTAACCAATCAG 4956
Db 3408 TTTCCAGGGTCCCGCAGGACTGAATGACCCCTGTGCTTATTGTAACTAACCAATCAG 3467
Qy 4957 TTCGGCTTCTCGCTTCTGTTCGGGGCTTCTGCTCCCGAGCTCAATAAAGAGCCCAAA 5016
Db 3468 TTCGGCTTCTCGCTTCTGTTCGGGGCTTCTGCTCCCGAGCTCAATAAAGAGCCCAAA 3527
Qy 5017 CCCCTCACTCGGGGCCAGTCTCTCCGATTGACTGAGTCGCGCGGTACCCGTTATCCA 5076
Db 3528 CCCCTCACTCGGGGCCAGTCTCTCCGATTGACTGAGTCGCGCGGTACCCGTTATCCA 3587
Qy 5077 ATAAACCTCTTCGAGTTGCATCCGACTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5136
Db 3588 ATAAACCTCTTCGAGTTGCATCCGACTTGTGGTCTGCTGCTGCTGCTGCTGCTGCTGCT 3647
Qy 5137 CTGAGTGATTGACTACCCGTCAGCGGGGGTCTTTTCATTTGG 5177
Db 3648 CTGAGTGATTGACTACCCGTCAGCGGGGGTCTTTTCATTTGG 3688

RESULT 11

US-10-987-388-36

; Sequence 36, Application US/10987388

; Publication No. US2005015871A1

; GENERAL INFORMATION:

; APPLICANT: Stratagene California

; APPLICANT: Braman, Jeffrey

; APPLICANT: Carstens, Carsten-Peter

; APPLICANT: Novoradovskaya, Natalia

; APPLICANT: Bagga, Rajesh

; APPLICANT: Baschore, Lee Scott

; TITLE OF INVENTION: Compositions and Methods for Protein Isolation

; FILE REFERENCE: 25436/2465

; CURRENT APPLICATION NUMBER: US/10/987,388

; CURRENT FILING DATE: 2004-11-12

; PRIOR APPLICATION NUMBER: US 10/712,137

; PRIOR FILING DATE: 2003-11-13

; NUMBER OF SEQ ID NOS: 36

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 36

; LENGTH: 6825

; TYPE: DNA

; ORGANISM: Artificial sequence

; FEATURE:

; OTHER INFORMATION: Retroviral vector pFB-CTAP-nso

US-10-987-388-36

Query Match 28.3%; Score 1463; DB 9; Length 6825;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1463; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 3715 GGATCGTTTCGATGATTGAACAGATGGATTGACGCGAGTCTCCGGCCGCTTGGGTG 3774
Db 3005 GGATCGTTTCGATGATTGAACAGATGGATTGACGCGAGTCTCCGGCCGCTTGGGTG 3064
Qy 3775 GAGAGGCTATTCCGCTATGACTGGGCAACAAGACAATCGGCTGCTCTGATGCGCCGCTG 3834
Db 3065 GAGAGGCTATTCCGCTATGACTGGGCAACAAGACAATCGGCTGCTCTGATGCGCCGCTG 3124
Qy 3835 TTCGGCTGTGACGCGAGGGGGCGCGTATCGTGGTGGCCACGACCTGTCGGGTGCC 3894
Db 3125 TTCGGCTGTGACGCGAGGGGGCGCGTATCGTGGTGGCCACGACCTGTCGGGTGCC 3184
Qy 3895 CTGAATGAATGACGAGGAGCGAGCGGCTATCGTGGTGGCCACGAGCGGGGTTCCT 3954
Db 3185 CTGAATGAATGACGAGGAGCGAGCGGCTATCGTGGTGGCCACGAGCGGGGTTCCT 3244
Qy 3955 TCGCGAGCTGTGCTCGAGTGTGTCACTGAAGCGGGAAGGAGCTGGCTGCTATTGGCGAA 4014
Db 3245 TCGCGAGCTGTGCTCGAGTGTGTCACTGAAGCGGGAAGGAGCTGGCTGCTATTGGCGAA 3304

QY 5095 GCATCCGACTTGTGCTCGCTGTTCTTGGAGGGTCTCTCTGAGTGATTGACTACCC 5154
Db |||||||
QY 4385 GCATCCGACTTGTGCTCGCTGTTCTTGGAGGGTCTCTCTGAGTGATTGACTACCC 4444
Db |||||||
QY 5155 GTCAGCGGGGCTCTTTCATTGG 5177
Db |||||||
QY 4445 GTCAGCGGGGCTTTCATTGG 4467
Db |||||||

RESULT 12

US-10-987-388-35
; Sequence 35, Application US/10987388
; Publication No. US2005015871A1
; GENERAL INFORMATION:
; APPLICANT: Stratagene California
; APPLICANT: Brame, Jeffrey
; APPLICANT: Carstens, Carsten-Peter
; APPLICANT: Novorodovskaya, Natalia
; APPLICANT: Bagga, Rajesh
; APPLICANT: Basenore, Lee Scott
; TITLE OF INVENTION: Compositions and Methods for Protein Isolation
; FILE REFERENCE: 25436/2465
; CURRENT APPLICATION NUMBER: US/10/987,388
; PRIORITY FILING DATE: 2004-11-12
; PRIOR APPLICATION NUMBER: US 10/712,137
; PRIOR FILING DATE: 2003-11-13
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 35
; LENGTH: 6851
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Retroviral vector pFB-NTAP-neo
US-10-987-388-35

Query Match 28.3%; Score 1463; DB 9; Length 6851;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1463; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 GGATCGTTTCCGATGATGAACAGATGGATTGACGAGGTCTCCGGCGCTGGGTG 3774
Db |||||||
QY 3031 GGATCGTTTCCGATGATGAACAGATGGATTGACGAGGTCTCCGGCGCTGGGTG 3090
Db |||||||
QY 3775 GAGAGGCTATTTCGGCTATGATCGGCAACAGACAAATCGGCTGCTGATGCGCGGTG 3834
Db |||||||
QY 3091 GAGAGGCTATTTCGGCTATGATCGGCAACAGACAAATCGGCTGCTGATGCGCGGTG 3150
Db |||||||
QY 3835 TTCGGCTGTGAGCGAGGGCGCGCGCTCTTTTGTCAAGACGACCTGTCGGTGCC 3894
Db |||||||
QY 3151 TTCGGCTGTGAGCGAGGGCGCGCGCTCTTTTGTCAAGACGACCTGTCGGTGCC 3210
Db |||||||
QY 3895 CTGAATGAATCGAGGAGGAGCGCGGTATGCTGCTGCTGCTGCTGCTGCTGCTGCT 3954
Db |||||||
QY 3211 CTGAATGAATCGAGGAGGAGCGCGGTATGCTGCTGCTGCTGCTGCTGCTGCT 3270
Db |||||||
QY 3955 TCGCGAGCTGTGCTGAGCTGTGCTCAAGCGGAGGAGCTGCTGCTATTTGGCGAA 4014
Db |||||||
QY 3271 TCGCGAGCTGTGCTGAGCTGTGCTCAAGCGGAGGAGCTGCTGCTATTTGGCGAA 3330
Db |||||||
QY 4015 GTCCGGGCGAGGATCTCTGCTATCTCACTTGTCTGCTGCGGAGAAATATCATG 4074
Db |||||||
QY 3331 GTCCGGGCGAGGATCTCTGCTATCTCACTTGTCTGCTGCGGAGAAATATCATG 3390
Db |||||||
QY 4075 GCTGATGAATCGGCGGCTGCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4134
Db |||||||
QY 3391 GCTGATGAATCGGCGGCTGCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3450
Db |||||||
QY 4135 GCGAAACATCGATCGAGGAGCGTACTCGGATGGAAGCGGTCTTGTGATCAGGAT 4194
Db |||||||
QY 3451 GCGAAACATCGATCGAGGAGCGTACTCGGATGGAAGCGGTCTTGTGATCAGGAT 3510
Db |||||||
QY 4195 GATCTGGACGAGAGCATCAGGGGCTCGCGCCAGCGAATGTTTCGCGAGGCTCAAGGCG 4254
Db |||||||

RESULT 13

US-10-359-397-1

; Sequence 1, Application US/10359397

; Publication No. US20040071673A1

; GENERAL INFORMATION:

APPLICANT: University of Southern California
; TITLE OF INVENTION: Construction and Use of Genes Encoding Pathogenic Epitopes for Treatment of Autoimmune Disease
; FILE REFERENCE: 2013761-7030803001
; CURRENT APPLICATION NUMBER: US/10/359,397
; CURRENT FILING DATE: 2003-02-05
; PRIOR APPLICATION NUMBER: 10/098,035
; PRIOR FILING DATE: 2002-03-14
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 1
; LENGTH: 5856
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Sal/hind drag of pSVNA into Sal/Hind cut G1
US-10-359-397-1

Query Match 28.1%; Score 1455.2; DB 7; Length 5856;
Best Local Similarity 98.4%; Pred. No. 0;
Matches 1488; Conservative 0; Mismatches 3; Indels 21; Gaps 1;

Qy 3687 ATCAAGATCTGATCAAGACAGCAGGATGAGGATCGTTTCGCAATGATTGAACAAGATGATT 3746
Db |||||
Qy 3747 GCAGCAGGTTCTCGGCGGCTTGGTGGAGAGGCTATTTCGGCTATGACTGGGCAACA 3806
Db |||||
Qy 1904 GCAGCAGGTTCTCGGCGGCTTGGTGGAGAGGCTATTTCGGCTATGACTGGGCAACA 1963
Db |||||
Qy 3807 GACAAATGGCTGCTGATGCGCGCTTCCGGCTGTCAGCGAGGGGCGCGGTTCT 3866
Db |||||
Qy 1964 GACAAATGGCTGCTGATGCGCGCTTCCGGCTGTCAGCGAGGGGCGCGGTTCT 2023
Db |||||
Qy 3867 TTTTGTCAAGACCGACCTGTCCGGTGCCTGAAATGAATCGCAGGACGAGGCGCGGCT 3926
Db |||||
Qy 2024 TTTTGTCAAGACCGACCTGTCCGGTGCCTGAAATGAATCGCAGGACGAGGCGCGGCT 2083
Db |||||
Qy 3927 ATCGTGGCTGGCCAGCAGCGGCGTTCTTGGCGAGCTGTGTCTGACGTTGTCACCTGAAGC 3986
Db |||||
Qy 2084 ATCGTGGCTGGCCAGCAGCGGCGTTCTTGGCGAGCTGTGTCTGACGTTGTCACCTGAAGC 2143
Db |||||
Qy 3987 GGGAAAGGACCTGGCTGCTATTGGGCGAAGTCCCGGGCAGGATCTCTGTCTCATCTCACCT 4046
Db |||||
Qy 2144 GGGAAAGGACCTGGCTGCTATTGGGCGAAGTCCCGGGCAGGATCTCTGTCTCATCTCACCT 2203
Db |||||
Qy 4047 TGCTCTCCGAGAAAGTATCATCATGGCTGATGCAATGCGGCGGCTGCTATACGCTTGA 4106
Db |||||
Qy 2204 TGCTCTCCGAGAAAGTATCATCATGGCTGATGCAATGCGGCGGCTGCTATACGCTTGA 2263
Db |||||
Qy 4107 TCCGGCTACCTGCCATTCGACCAACCAAGCAATCGCATCGAGCGAGCACGTAATCG 4166
Db |||||
Qy 2264 TCCGGCTACCTGCCATTCGACCAACCAAGCAATCGCATCGAGCGAGCACGTAATCG 2323
Db |||||
Qy 4167 GATGAAAGCGGCTGTTCGATCAGGATGATCTGGAAGAGCATCAGGGGCTCGGCGC 4236
Db |||||
Qy 2324 GATGAAAGCGGCTGTTCGATCAGGATGATCTGGAAGAGCATCAGGGGCTCGGCGC 2383
Db |||||
Qy 4227 AGCCGAACTGTTCGCGAGGCTCAAGGCGCGCATGCCGACGCGGAGGATCTCGTCGTGAC 4286
Db |||||
Qy 2384 AGCCGAACTGTTCGCGAGGCTCAAGGCGCGCATGCCGACGCGGAGGATCTCGTCGTGAC 2443
Db |||||
Qy 4287 CCATGGCGATGCTGCTTGGCGAATATCATGGTGGAAATGGCGGCTTTCTGGAATCAT 4346
Db |||||
Qy 2444 CCATGGCGATGCTGCTTGGCGAATATCATGGTGGAAATGGCGGCTTTCTGGAATCAT 2503
Db |||||
Qy 4347 CGACTGTGGCGGCTGGGTGGCGGCGCTATCAGGACATAGCTTGGCTACCCGTA 4406
Db |||||
Qy 2504 CGACTGTGGCGGCTGGGTGGCGGCGCTATCAGGACATAGCTTGGCTACCCGTA 2563
Db |||||
Qy 4407 TATTGCTCAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCTCGTGTCTTACGGTATCGC 4466
Db |||||
Qy 2564 TATTGCTCAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCTCGTGTCTTACGGTATCGC 2623
Db |||||

Qy 4467 CGTCCCGATTTCGAGCGCATCGCTTCTATCGCTTCTTTGACGAGTTCTTCTGAGCGGG 4526
Db |||||
Qy 2624 CGTCCCGATTTCGAGCGCATCGCTTCTATAGCTTCTTTGACGAGTTCTTCTGAGCGGG 2683
Db |||||
Qy 4527 ACTCTGGGGT-----TCGATAAAATAAAAGATTTTATTAGTCT 4565
Db |||||
Qy 2684 ACTCTGGGGTTCGTCGAGAGCTTGGGCCCATCGATAAAATAAAAGATTTTATTAGTCT 2743
Db |||||
Qy 4566 CAGAAAAAGGGGGGAATGAAAGACCCACCTGTAGGTTTGGCAAGCTAGCTTAAAGTAAC 4625
Db |||||
Qy 2744 CAGAAAAAGGGGGGAATGAAAGACCCACCTGTAGGTTTGGCAAGCTAGCTTAAAGTAAC 2803
Db |||||
Qy 4626 GCCATTTTCAAGGCAATGGAATAATACATAACTAGAGAAATAGAGAGTTTCAGATCAAGGTC 4685
Db |||||
Qy 2804 GCCATTTTCAAGGCAATGGAATAATACATAACTAGAGAAATAGAGAAATTCAGATCAAGGTC 2863
Db |||||
Qy 4686 AGGAAACAGATGGAACAGCTGAATATGAGGCCAAACAGGATATCTGTGTAAGCAGTTCCCTG 4745
Db |||||
Qy 2864 AGGAAACAGATGGAACAGCTGAATATGAGGCCAAACAGGATATCTGTGTAAGCAGTTCCCTG 2923
Db |||||
Qy 4746 CCCCGGCTCAGGGCCAAAGAACAGATGGAACAGCTGAATATGAGGCCAAACAGGATATCTGT 4805
Db |||||
Qy 2924 CCCCGGCTCAGGGCCAAAGAACAGATGGAACAGCTGAATATGAGGCCAAACAGGATATCTGT 2983
Db |||||
Qy 4806 GGTAAAGCAGTTCTGCTGCGGCTCAGGGCCAAAGAACAGATGTTCCCGAGATGCCGTCAG 4865
Db |||||
Qy 2984 GGTAAAGCAGTTCTGCTGCGGCTCAGGGCCAAAGAACAGATGTTCCCGAGATGCCGTCAG 3043
Db |||||
Qy 4866 CCCTCAGCAGTTCTTAGAGAACCATCAGATGTTTCCAGGGTGCCCGCAAGGACCTGAAATG 4925
Db |||||
Qy 3044 CCCTCAGCAGTTCTTAGAGAACCATCAGATGTTTCCAGGGTGCCCGCAAGGACCTGAAATG 3103
Db |||||
Qy 4926 ACCTGTGCTTATTGTAACCAATCAGTTGCTTCTGCTTCTGCTTCTGTCGCGGCTTC 4985
Db |||||
Qy 3104 ACCTGTGCTTATTGTAACCAATCAGTTGCTTCTGCTTCTGCTTCTGTCGCGGCTTC 3163
Db |||||
Qy 4986 TGCTCCCGAGCTCAATAAAGAGCCCAACCCCTCACCTCGGGGCGCAGTCCCTCCGAT 5045
Db |||||
Qy 3164 TGCTCCCGAGCTCAATAAAGAGCCCAACCCCTCACCTCGGGGCGCAGTCCCTCCGAT 3223
Db |||||
Qy 5046 TGACTGAGTCCCGGCTACCCGCTGATCAATAAACCCCTCTTGCACTGTCATCCGACTT 5105
Db |||||
Qy 3224 TGACTGAGTCCCGGCTACCCGCTGATCAATAAACCCCTCTTGCACTGTCATCCGACTT 3283
Db |||||
Qy 5106 GTGGTCTCGCTTCTTGGGAGGCTCTCTCTGAGTGATGACTACCGCTCAGCGGGG 5165
Db |||||
Qy 3284 GTGGTCTCGCTTCTTGGGAGGCTCTCTCTGAGTGATGACTACCGCTCAGCGGGG 3343
Db |||||
Qy 5166 TCTTTTCATTGG 5177
Db |||||
Qy 3344 TCTTTTCATTGG 3355
Db |||||

RESULT 14
US-10-098-035-3
; Sequence 3, Application US/10098035
; Publication No. US20020141983A1
; GENERAL INFORMATION:
; APPLICANT: University of Southern California
; APPLICANT: Weiner, Leslie P.
; APPLICANT: McMillan, Minnie
; TITLE OF INVENTION: CONSTRUCTION AND USE OF GENES ENCODING
; TITLE OF INVENTION: PATHOGENIC EPITOPES FOR TREATMENT OF AUTOIMMUNE DISEASE
; FILE REFERENCE: 13761-703-00 US
; CURRENT APPLICATION NUMBER: US/10/098,035
; CURRENT FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: US/08/654,737
; PRIOR FILING DATE: 1996-05-29
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 5865

Db 2639 ATTGAACAAGATGGATTGTCAGCGAGGTTCTCCGGCGCTTGCGTGAGAGGCTATTTCGGC 2698
Qy 3790 TATGACTGGGCACACAGACAAATCGCTGCTGATGCGCGTGTTCGGGCTGTACGG 3849
Db 2699 TATGACTGGGCACACAGACAAATCGCTGCTGATGCGCGTGTTCGGGCTGTACGG 2758
Qy 3850 CAGGGGCGCCCGGTTCTTTTGTCAAGACCGACCTGTCCGGTGCCCTGAATGAATGCGAG 3909
Db 2759 CAGGGGCGCCCGGTTCTTTTGTCAAGACCGACCTGTCCGGTGCCCTGAATGAATGCGAG 2818
Qy 3910 GACGAGGACGCGCGCTATCGTGGCTGCGCAGCAGCGCGTTCCTTGCGCAGCTGTGCTC 3969
Db 2819 GACGAGGACGCGCGCTATCGTGGCTGCGCAGCAGCGCGTTCCTTGCGCAGCTGTGCTC 2878
Qy 3970 GACGTTGTCACTGAAGCGGGAAGGACTGGCTGTATTGGGGAAGTCCCGGGCAGGAT 4029
Db 2879 GACGTTGTCACTGAAGCGGGAAGGACTGGCTGTATTGGGGAAGTCCCGGGCAGGAT 2938
Qy 4030 CTCTCTGTCACTCACTTGTCTCTCCCGAAGAAATATCCATCATGGCTGATGCAATGGG 4089
Db 2939 CTCTCTGTCACTCACTTGTCTCTCCCGAAGAAATATCCATCATGGCTGATGCAATGGG 2998
Qy 4090 CGGCTGCAATAGCTTGTATCCGGCTACTGCGCATTCGACCAAGCGAAACATGCGATC 4149
Db 2999 CGGCTGCAATAGCTTGTATCCGGCTACTGCGCATTCGACCAAGCGAAACATGCGATC 3058
Qy 4150 GACGAGCAGCTACTCGATGGAAGCGGCTTGTGCGATCAGGATGATCTCGACGAAGAG 4209
Db 3059 GACGAGCAGCTACTCGATGGAAGCGGCTTGTGCGATCAGGATGATCTCGACGAAGAG 3118
Qy 4210 CATCAGGGGCTCGCCGACCGGAACTGTTGCGCAGGCTCAAGGGCGCATGCCGACGGC 4269
Db 3119 CATCAGGGGCTCGCCGACCGGAACTGTTGCGCAGGCTCAAGGGCGCATGCCGACGGC 3178
Qy 4270 GAGGATCTCGTGTGACCCATGGCGATGCTGCTTGCAGAAATATCATGGTGGAAAAATGGC 4329
Db 3179 GAGGATCTCGTGTGACCCATGGCGATGCTGCTTGCAGAAATATCATGGTGGAAAAATGGC 3238
Qy 4330 CGCTTTTCTGGATTCACTGACTGTGGCCGGTGGGTGTGGCGGACCGCTATCAGGACATA 4389
Db 3239 CGCTTTTCTGGATTCACTGACTGTGGCCGGTGGGTGTGGCGGACCGCTATCAGGACATA 3298
Qy 4390 GGGTTGGCTACCGGTGATATTGCTGAAGACTTGGCGCGAATGGGCTGACCGTTCCTC 4449
Db 3299 GGGTTGGCTACCGGTGATATTGCTGAAGACTTGGCGCGAATGGGCTGACCGTTCCTC 3358
Qy 4450 GTGCTTTACGGTATCGCCGCTCCGATTTCGACGCGCATCGCTTCTATCGCTTCTTGAC 4509
Db 3359 GTGCTTTACGGTATCGCCGCTCCGATTTCGACGCGCATCGCTTCTATCGCTTCTTGAC 3418
Qy 4510 GAGTTCTTCTGAGCGGACTCTGGGGTTCGATAAAATAAAGATTTTATTAGTCTCCAG 4569
Db 3419 GAGTTCTTCTGAGCGGACTCTGGGGTTCGATAAAATAAAGATTTTATTAGTCTCCAG 3478
Qy 4570 AAAAAGGGGGAAATCAAAAGACCCACCTGTAGTTTGGCAAGCTAGCTTAAGTAACGCCA 4629
Db 3479 AAAAAGGGGGAAATCAAAAGACCCACCTGTAGTTTGGCAAGCTAGCTTAAGTAACGCCA 3538
Qy 4630 TTTTGCAGGCAATGGAATAATACATACTGAGAATAGAGAGTTTCAGATCAAGGTCAAGGA 4689
Db 3539 TTTTGCAGGCAATGGAATAATACATACTGAGAATAGAGAGTTTCAGATCAAGGTCAAGGA 3598
Qy 4690 ACAGATGGAAATGCTGAATATGGGCAAAACAGGATATCTGTGGTAAGCAGTTCTTGCCCC 4749
Db 3599 ACAGATGGAAATGCTGAATATGGGCAAAACAGGATATCTGTGGTAAGCAGTTCTTGCCCC 3658
Qy 4750 GCTCAGGCGCAAGAACAGATGGAAACAGTGAATATGGGCAAAACAGGATATCTGTGGTA 4809
Db 3659 GCTCAGGCGCAAGAACAGATGGAAACAGTGAATATGGGCAAAACAGGATATCTGTGGTA 3718
Qy 4810 AGCAGTTCTGCCCCGGCTCAGGGGCAAGAAACAGATGGTCCCGCAGATCGCGTCCAGCCCT 4869
Db 3719 AGCAGTTCTGCCCCGGCTCAGGGGCAAGAAACAGATGGTCCCGCAGATCGCGTCCAGCCCT 3778

Qy 4870 CAGCAGTTTCTAGAGAAACCATCAGATGTTTCCAGGGTGCCCCAAGGACCTGAAATGACCC 4929
Db 3779 CAGCAGTTTCTAGAGAAACCATCAGATGTTTCCAGGGTGCCCCAAGGACCTGAAATGACCC 3838
Qy 4930 TGTGCTTTATTGAACTAAACCAATCAGTTGCTTCTCGCTTCTGTTGCGGCGCTTCTGCT 4989
Db 3839 TGTGCTTTATTGAACTAAACCAATCAGTTGCTTCTCGCTTCTGTTGCGGCGCTTCTGCT 3898
Qy 4990 CCGGAGGCTCAATAAAAGAGCCCAAAACCCCTCACTCGGGGCGCAGTCTCCGATTGAC 5049
Db 3899 CCGGAGGCTCAATAAAAGAGCCCAAAACCCCTCACTCGGGGCGCAGTCTCCGATTGAC 3958
Qy 5050 TGAGTCGCGCGGGTACCCGCTGATCCCAATAAACCTCTTGCAATTGCATCCGACTTGTGG 5109
Db 3959 TGAGTCGCGCGGGTACCCGCTGATCCCAATAAACCTCTTGCAATTGCATCCGACTTGTGG 4018
Qy 5110 TCTGCTGTTCTCTTGGGAGGGTCTCTCTGAGTGAATTGACTACCCGTCAGCGGGGCTT 5169
Db 4019 TCTGCTGTTCTCTTGGGAGGGTCTCTCTGAGTGAATTGACTACCCGTCAGCGGGGCTT 4078
Qy 5170 TCATTGG 5177
Db 4079 TCATTGG 4086

Search completed: February 27, 2006, 10:26:20
Job time : 3831 secs


```
Db CCTCTGTAAGTTGGGAAGCCCTGCARAGTAACATGATGGCTTCTTTCGGCCCAAGGA 3569
Qy TCTGATGCGCAGGGGATCAAGATCTGATCAAGAGACAGAGTAGAGATCGTTTCGCATGA 3730
Db TCTGATGCGCAGGGGATCAAGATCTGATCAAGAGACAGAGTAGAGATCGTTTCGCATGA 3509
Qy TTGAACAGATGAGTTGACCCAGGTTCTCCGGCCGCTTGGGTGGAGAGGCTATTTCGGCT 3790
Db TTGAACAAGATGAGTTGACCCAGGTTCTCCGGCCGCTTGGGTGGAGAGGCTATTTCGGCT 3449
Qy ATGACTGGGCAACACAGACAATCGGCTGCTCTGATGCCCGCTGTTCCGGCTGTTCAGCGC 3850
Db ATGACTGGGCAACACAGACAATCGGCTGCTCTGATGCCCGCTGTTCCGGCTGTTCAGCGC 3389
Qy AGGGGCGCCCGGTTCTTTTGTCAAGACCGACCTGTCTCGGCTGCTGATGAATGAATGCAAG 3910
Db AGGGGCGCCCGGTTCTTTTGTCAAGACCGACCTGTCTCGGCTGCTGATGAATGAATGCAAG 3329
Qy ACGAGGACGCGGCTATCGTGCTGCGTGGCCACGACGCGGCTTCTTTCGCGAGCTGTGCTCG 3970
Db ACGAGGACGCGGCTATCGTGCTGCGTGGCCACGACGCGGCTTCTTTCGCGAGCTGTGCTCG 3269
Qy ACGTTGTCACTGAACGCGGAAGGACTGGCTGCTATTGGGCGAAGTGGCGGGCAGGATC 4030
Db ACGTTGTCACTGAACGCGGAAGGACTGGCTGCTATTGGGCGAAGTGGCGGGCAGGATC 3209
Qy TCTGTCTATCTACCTTCTGCTCGGAGAAAGTATCCATCATGCTGATGCAATGCGGC 4090
Db TCTGTCTATCTACCTTCTGCTCGGAGAAAGTATCCATCATGCTGATGCAATGCGGC 3149
Qy GGCTGCATACGCTTGTATCGGCTACCTGCCCAATTCGACCAACGCGAAACATGCGCATCG 4150
Db GGCTGCATACGCTTGTATCGGCTACCTGCCCAATTCGACCAACGCGAAACATGCGCATCG 3089
Qy AGCGAGACGCTACTCGGATGGAAGCGGCTTGTGTCGATCAGGATGATCTGACCAAGAGC 4210
Db AGCGAGACGCTACTCGGATGGAAGCGGCTTGTGTCGATCAGGATGATCTGACCAAGAGC 3029
Qy ATCAGGGGCTCGCGCAGCCGAACTGTTCGACAGGCTCAAGCGCGCATGCCGACGCGC 4270
Db ATCAGGGGCTCGCGCAGCCGAACTGTTCGACAGGCTCAAGCGCGCATGCCGACGCGC 2969
Qy AGGATCTGCTGATGCCATGCGATGCTGCTGCGGAAATATCATGCTGGAATATGCCC 4330
Db AGGATCTGCTGATGCCATGCGATGCTGCTGCGGAAATATCATGCTGGAATATGCCC 2909
Qy GCTTTTCTGGATTTCATCGACTGTGGCGGCTGGGTGTGGCGACCGCTATCAGGACATAG 4390
Db GCTTTTCTGGATTTCATCGACTGTGGCGGCTGGGTGTGGCGACCGCTATCAGGACATAG 2849
Qy CGTTGGCTACCGTGATATGCTGAAGAGCTTGGCGGCAATGGGCTGACCGCTTCCTCG 4450
Db CGTTGGCTACCGTGATATGCTGAAGAGCTTGGCGGCAATGGGCTGACCGCTTCCTCG 2789
Qy TGCCTTACGATATGCGGCTCCGATTCGACGCGATCGCTTCTATCGCTTCTTGACG 4510
Db TGCCTTACGATATGCGGCTCCGATTCGACGCGATCGCTTCTATCGCTTCTTGACG 2729
Qy AGTTCTTCTGAGCGGACTCTGGGGTTTCGATA 4542
Db AGTTCTTCTGAGCGGACTCTGGGGTTTCGATA 2697
```

RESULT 2

US-11-233-119-16/c

; Sequence 16, Application US/11233119

; Publication No. US20060025331A1

; GENERAL INFORMATION:

; APPLICANT: HU, JING-SHAN et al,

; TITLE OF INVENTION: VASCULAR ENDOTHELIAL GROWTH FACTOR 2

; FILE REFERENCE: PF112P4D1

; CURRENT APPLICATION NUMBER: US/11/233,119

```
; CURRENT FILING DATE: 2005-09-23
; PRIOR APPLICATION NUMBER: 09/107,997
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: 09/042,105
; PRIOR FILING DATE: 1998-03-13
; PRIOR APPLICATION NUMBER: 09/999,811
; PRIOR FILING DATE: 1997-12-24
; PRIOR APPLICATION NUMBER: 08/465,968
; PRIOR FILING DATE: 1995-06-06
; PRIOR APPLICATION NUMBER: 08/207,550
; PRIOR FILING DATE: 1994-03-08
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 16
; LENGTH: 3974
; TYPE: DNA
; ORGANISM: Expression vector pHEA4-5
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(3974)
; OTHER INFORMATION: Expression vector pHEA4-5
; US-11-233-119-16
```

Query Match 22.6%; Score 1170.4; DB 12; Length 3974;

Best Local Similarity 99.9%; Pred. No. 8.7e-268;

Matches 1171; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy 3371 AAGTTCCAGCTGCCGCAAGCACTCAGGGCGCAAGGGCTGTCTAAAGAGCGGAACAGCT 3430
Db 3868 AAGTTCCAGCTGCCGCAAGCACTCAGGGCGCAAGGGCTGTCTAAAGAGCGGAACAGCT 3809
Qy 3431 AGAAGCGCACTCCGCAAGACGGTCTGACCCCGATGAATGTCTAGCTTCTTGGCTATCT 3490
Db 3808 AGAAGCGCACTCCGCAAGACGGTCTGACCCCGATGAATGTCTAGCTTCTTGGCTATCT 3749
Qy 3491 GGACAGCGGAAACCGCAAGCGCAAGAGAAAGAGAGAGTAGCTTGCAGTGGGCTTACATGGC 3550
Db 3748 GGACAGCGGAAACCGCAAGCGCAAGAGAGAAAGAGAGTAGCTTGCAGTGGGCTTACATGGC 3689
Qy 3551 GATAGCTAGACTGGGCGGTTTTATGGACAGCAAGCGAAATTCGAGCTGGGCGC 3610
Db 3688 GATAGCTAGACTGGGCGGTTTTATGGACAGCAAGCGAAATTCGAGCTGGGCGC 3629
Qy 3611 CCTCTGTAGGTTGGGAAGCCCTGCAAGATGAATCTGATGGCTTCTTTCGCGCAAGGA 3670
Db 3628 CCTCTGTAGGTTGGGAAGCCCTGCAAGATGAATCTGATGGCTTCTTTCGCGCAAGGA 3569
Qy 3671 TCTGATGCGCGCAGGGGATCAAGATCTGATCAAGAGACAGGATGAGGATCGTTTCGCATGA 3730
Db 3568 TCTGATGCGCGCAGGGGATCAAGATCTGATCAAGAGACAGGATGAGGATCGTTTCGCATGA 3509
Qy 3731 TTGAACAAGATGAGTTGCAAGCGGCTTCTCGGCGGCTTGGGTGGAGAGGCTATTTCGGCT 3790
Db 3508 TTGAACAAGATGAGTTGCAAGCGGCTTCTCGGCGGCTTGGGTGGAGAGGCTATTTCGGCT 3449
Qy 3791 ATGACTGGGCAACACAGACAATCGGCTGCTGATGCCCGCTGTTCCGGCTGTTCAGCGC 3850
Db 3448 ATGACTGGGCAACACAGACAATCGGCTGCTGATGCCCGCTGTTCCGGCTGTTCAGCGC 3389
Qy 3851 AGGGGCGCCCGGTTCTTTTGTCAAGACCGACCTGTCTCGGTCGCCCTGAATGAACTGCAGG 3910
Db 3388 AGGGGCGCCCGGTTCTTTTGTCAAGACCGACCTGTCTCGGTCGCCCTGAATGAACTGCAGG 3329
Qy 3911 ACGAGGACGCGGCTATCGTGCTGGCCACAGACGCGGCTTCTTTCGCGAGCTGTGCTCG 3970
Db 3328 ACGAGGACGCGGCTATCGTGCTGGCCACAGACGCGGCTTCTTTCGCGAGCTGTGCTCG 3269
Qy 3971 ACGTTGTCACTGAAGCGGGAAGGACTGGCTGCTATTGGGCGAAGTGGCGGGCAGGATC 4030
Db 3268 ACGTTGTCACTGAAGCGGGAAGGACTGGCTGCTATTGGGCGAAGTGGCGGGCAGGATC 3209
Qy 4031 TCCTGTCTATCTACCTTCTGCTCGGAGAAAGTATCCATCATGCTGATGCAATGCGGC 4090
```


Db 3208 TCCTGTGATCTCACCTTCTCTCCGAGAAAGTATCCATCATGCTGATGCAATGGGC 3149
Qy 4091 GGTGTCATACGCTTGTATCCGCTACCTCCCAATTCGACCAACGCAAAATCGCATCG 4150
Db 3148 GGTGTCATACGCTTGTATCCGCTACCTCCCAATTCGACCAACGCAAAATCGCATCG 3089
Qy 4151 AGCGAGCAGCTACTCGGATGGAAGCCGCTCTTGTGATCAGGATGATCTGACGAAGGC 4210
Db 3088 AGCGAGCAGCTACTCGGATGGAAGCCGCTCTTGTGATCAGGATGATCTGACGAAGGC 3029
Qy 4211 ATCAGGGCTCGCGCAGCCGAATCTGCGAGGCTCAAGCGCGCATGCCGACGCGC 4270
Db 3028 ATCAGGGCTCGCGCAGCCGAATCTGCGAGGCTCAAGCGCGCATGCCGACGCGC 2969
Qy 4271 AGGATCTCGTGTGACCCATCGCGATCGCTCTGCGGAATATCATGTGGAAATGGCC 4330
Db 2968 AGGATCTCGTGTGACCCATCGCGATCGCTCTGCGGAATATCATGTGGAAATGGCC 2909
Qy 4331 GCTTTCTGGAATTCATGACATGTGCGCGCTGGGTGTGGCGGATGATCAGGACATAG 4390
Db 2908 GCTTTCTGGAATTCATGACATGTGCGCGCTGGGTGTGGCGGATGATCAGGACATAG 2849
Qy 4391 GCTTGGCTTACCCGTGATATCTGAAGAGCTTGGCGGGAATGGCTGACCGCTTCTCG 4450
Db 2848 GCTTGGCTTACCCGTGATATCTGAAGAGCTTGGCGGGAATGGCTGACCGCTTCTCG 2789
Qy 4451 TGCTTTACGCTATCGCGCTCCGATTCGAGCGCATCGCTTCTATCGCCTTCTTGACG 4510
Db 2788 TGCTTTACGCTATCGCGCTCCGATTCGAGCGCATCGCTTCTATCGCCTTCTTGACG 2729
Qy 4511 AGTTCTTCTGAGCGGACTCTGGGTTTCGATA 4542
Db 2728 AGTTCTTCTGAGCGGACTCTGGGTTTCGAAA 2697

RESULT 3

US-10-948-344-2
; Sequence 2, Application US/10948344
; Publication No. US20060013805A1
; GENERAL INFORMATION:
; APPLICANT: Hebbel, R.P.
; APPLICANT: Lin, Y.
; APPLICANT: Lollar, J.S.
; TITLE OF INVENTION: Transgenic circulating endothelial cells
; FILE REFERENCE: 600.449US1
; CURRENT APPLICATION NUMBER: US/10/948,344
; CURRENT FILING DATE: 2004-09-23
; PRIOR APPLICATION NUMBER: US/09/865,022
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: PCT/US99/28033
; PRIOR FILING DATE: 1999-11-24
; PRIOR APPLICATION NUMBER: US 60/109,687
; PRIOR FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 12445
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: The DNA sequence of HSQRENeo.
US-10-948-344-2

Query Match 22.6%; Score 1170.4; DB 7; Length 12445;
Best Local Similarity 99.9%; Pred.No.1.5e-267;
Matches 1171; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 3371 AAGCTTCAAGCTGCGGCAAGCACTCAGGCGCAAGGCTGTAAAGGAAGCGGACACGT 3430
Db 7480 AAGCTTCAAGCTGCGGCAAGCACTCAGGCGCAAGGCTGTAAAGGAAGCGGACACGT 7539
Qy 3431 AGAAGCCAGTCCGAGAAACGGTGTGACCCCGGATGAATGTACGCTATCTGGGCTATCT 3490

Db 7540 AGAAGCCAGTCCGAGAAACGGTGTGACCCCGGATGAATGTACGCTATCTGGGCTATCT 7599
Qy 3491 GGACAGGGAAGAAACGCAAGCGCAAGAGAAAGCAGTAGCTTTGAGTGGGCTTACATGGC 3550
Db 7600 GGACAGGGAAGAAACGCAAGCGCAAGAGAAAGCAGTAGCTTTGAGTGGGCTTACATGGC 7659
Qy 3551 GATAGCTAGACTGGGCGGTTTTATGGACAGCAAGCGAAACCGGAATTCGACGCTGGGGCGC 3610
Db 7660 GATAGCTAGACTGGGCGGTTTTATGGACAGCAAGCGAAACCGGAATTCGACGCTGGGGCGC 7719
Qy 3611 CCTCTGTGTAGGTTGGGAAGCCCTGCAAAAGTAAACTGATGCTTCTTTCGCCCAAGGA 3670
Db 7720 CCTCTGTGTAGGTTGGGAAGCCCTGCAAAAGTAAACTGATGCTTCTTTCGCCCAAGGA 7779
Qy 3671 TCTGATCGCGCAGGGGATCAAGATCTGATCAAGAGACAGAGATGAGGATCGTTTCGCAATGA 3730
Db 7780 TCTGATCGCGCAGGGGATCAAGATCTGATCAAGAGACAGAGATGAGGATCGTTTCGCAATGA 7839
Qy 3731 TTGAACAAGATGGATTTGACAGCGAGTTCTCGGCGGCTTTGGGTGGAGAGGCTATTCGGCT 3790
Db 7840 TTGAACAAGATGGATTTGACAGCGAGTTCTCGGCGGCTTTGGGTGGAGAGGCTATTCGGCT 7899
Qy 3791 ATGACTCGGCAACACAGCAATCGGCTGCTCTGATGCGCGCTTTCGGGTGGAGAGGCTATTCGGCT 3850
Db 7900 ATGACTCGGCAACACAGCAATCGGCTGCTCTGATGCGCGCTTTCGGGTGGAGAGGCTATTCGGCT 7959
Qy 3851 AGGGCGCGCGGCTTTCTTTTGTCAAGACCGACCTGTGCGGTGCGCTGAATGAATCGATCG 3910
Db 7960 AGGGCGCGCGGCTTTCTTTTGTCAAGACCGACCTGTGCGGTGCGCTGAATGAATCGATCG 8019
Qy 3911 ACAGGACAGCGCGCTATCGTGGCTTGGCCACAGCGGGGTTCTTTCGCGACGCTGTCTCG 3970
Db 8020 ACAGGACAGCGCGCTATCGTGGCTTGGCCACAGCGGGGTTCTTTCGCGACGCTGTCTCG 8079
Qy 3971 ACCTTGTCACTGAAGCGGAGGACTGGCTGCTATTGGGCGAAGTGCCTGGGCGAGGATC 4030
Db 8080 ACCTTGTCACTGAAGCGGAGGACTGGCTGCTATTGGGCGAAGTGCCTGGGCGAGGATC 8139
Qy 4031 TCCTGTCTATCTCACTTTGCTCTCGCGAGAAAGTATCCATCATGCGCTGATGCAATTCGCGC 4090
Db 8140 TCCTGTCTATCTCACTTTGCTCTCGCGAGAAAGTATCCATCATGCGCTGATGCAATTCGCGC 8199
Qy 4091 GGCTGCATACGCTTGTATCGGCTACCTGCCCATTCGACCAAGCGGAAACATCGCATCG 4150
Db 8200 GGCTGCATACGCTTGTATCGGCTACCTGCCCATTCGACCAAGCGGAAACATCGCATCG 8259
Qy 4151 AGCGAGCAGCTACTCGGATGGAAGCCGCTCTTGTGATCAGGATGATCTGCAAGAGGC 4210
Db 8260 AGCGAGCAGCTACTCGGATGGAAGCCGCTCTTGTGATCAGGATGATCTGCAAGAGGC 8319
Qy 4211 ATCAGGGCTCGCGCAGCGCAACTGTTTCGCGAGGCTCAAGCGCGCATGCCGACGCGC 4270
Db 8320 ATCAGGGCTCGCGCAGCGCAACTGTTTCGCGAGGCTCAAGCGCGCATGCCGACGCGC 8379
Qy 4271 AGGATCTCGTGTGACCCATGGCGATCGCTGTTTCGCGAATATCATGTGAAATTCGCGC 4330
Db 8380 AGGATCTCGTGTGACCCATGGCGATCGCTGTTTCGCGAATATCATGTGAAATTCGCGC 8439
Qy 4331 GCTTTCTGGAATTCATGACATGCGCGCTGGGTGTGGCGGAGCGCTATCAGGACATAG 4390
Db 8440 GCTTTCTGGAATTCATGACATGCGCGCTGGGTGTGGCGGAGCGCTATCAGGACATAG 8499
Qy 4391 CGTTGGCTTACCCGCGGATTTGCTGAAGAGCTTGGCGGGAATGGGCTTACCGCTTCTCG 4450
Db 8500 CGTTGGCTTACCCGCGGATTTGCTGAAGAGCTTGGCGGGAATGGGCTTACCGCTTCTCG 8559
Qy 4451 TGCTTTACGCTATCGCGCTCCCGGATTCGAGCGCATCGCTTCTATCGCTTCTTGACG 4510
Db 8560 TGCTTTACGCTATCGCGCTCCCGGATTCGAGCGCATCGCTTCTATCGCTTCTTGACG 8619
Qy 4511 AGTTCTTCTGAGCGGAGCTCTGGGTTTCGATA 4542
Db 8620 AGTTCTTCTGAGCGGAGCTCTGGGTTTCGAAA 8651

RESULT 4

US-11-053-187-16/c
; Sequence 16, Application US/11053187
; Publication No. US20050282184A1
; GENERAL INFORMATION:
; APPLICANT: INVITROGEN CORPORATION
; APPLICANT: CHESNUT, Jonathan D.
; APPLICANT: SHUMAN, Stewart
; APPLICANT: MADDEN, Knut R.
; APPLICANT: HEYMAN, John A.
; APPLICANT: BENNETT, Robert P.
; TITLE OF INVENTION: METHODS AND REAGENTS FOR MOLECULAR CLONING
; FILE REFERENCE: INVIT1300-1
; CURRENT APPLICATION NUMBER: US/11/053,187
; CURRENT FILING DATE: 2005-02-07
; PRIOR FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US/09/935,280
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/226,563
; PRIOR FILING DATE: 2000-08-21
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 16
; LENGTH: 2290
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: puni/V5-His version A vector
US-11-053-187-16

Query Match 22.6%; Score 1167.8; DB 12; Length 2290;

Best Local Similarity 99.8%; Pred. No. 2.8e-267;
Matches 1169; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

| | | | |
|----|------|--|------|
| Qy | 3372 | AGCTTACGCTGCGCAGCACTCAGGCGCAAGGGCTGCTAAAGGAAGCGAACACGTA | 3431 |
| Db | 2290 | AGCTTACGCTGCGCAGCACTCAGGCGCAAGGGCTGCTAAAGGAAGCGAACACGTA | 2231 |
| Qy | 3432 | GAAGCCAGTCCGAGAAACGGTGTGACCCCGGATGAATCTCAGCTACTGGGCTATCTG | 3491 |
| Db | 2230 | GAAGCCAGTCCGAGAAACGGTGTGACCCCGGATGAATCTCAGCTACTGGGCTATCTG | 2171 |
| Qy | 3492 | GACAAAGGAAACCGAACGCAAGAGAAAGACAGCTAGCTTGCAGTGGGCTTACATGGCG | 3551 |
| Db | 2170 | GACAAAGGAAACCGAACGCAAGAGAAAGACAGCTAGCTTGCAGTGGGCTTACATGGCG | 2111 |
| Qy | 3552 | ATAGCTAGCTGGGCGGTTTATGACACAGCAAGCGAACCGGAATGGCCAGCTGGGGCGCC | 3611 |
| Db | 2110 | ATAGCTAGCTGGGCGGTTTATGACACAGCAAGCGAACCGGAATGGCCAGCTGGGGCGCC | 2051 |
| Qy | 3612 | CTCTGGTAAGGTTGGGAAGCCCTGCAAGTAACTGGATGGCTTTCTTGGCGCCAAGGAT | 3671 |
| Db | 2050 | CTCTGGTAAGGTTGGGAAGCCCTGCAAGTAACTGGATGGCTTTCTTGGCGCCAAGGAT | 1991 |
| Qy | 3672 | CTGATGGCGCAGGGATCAAGATCTGATCAAGACAGAGATGAGATCGTTTCGCATGAT | 3731 |
| Db | 1990 | CTGATGGCGCAGGGATCAAGATCTGATCAAGACAGAGATGAGATCGTTTCGCATGAT | 1931 |
| Qy | 3732 | TGAACAGATCGATTGACAGCAGGTTCTCCGGCGCTTGGGTGGAGAGGCTATTCCGGCTA | 3791 |
| Db | 1930 | TGAACAGATCGATTGACAGCAGGTTCTCCGGCGCTTGGGTGGAGAGGCTATTCCGGCTA | 1871 |
| Qy | 3792 | TGACTGGGCACACAGACAATCGGCTGCTCTGATGCGCGCGCTGTTCCGGCTGTCCAGGGCA | 3851 |
| Db | 1870 | TGACTGGGCACACAGACAATCGGCTGCTCTGATGCGCGCGCTGTTCCGGCTGTCCAGGGCA | 1811 |
| Qy | 3852 | GGGGCGCCGGTCTTTTGTCAAGACCGACTGTCCGGTCCCTGATGAATGAATGCGAGGA | 3911 |
| Db | 1810 | GGGGCGCCGGTCTTTTGTCAAGACCGACTGTCCGGTCCCTGATGAATGAATGCGAGGA | 1751 |
| Qy | 3912 | CGAGGACGCGGGCTATCGTGCTTGGCTGGCCACGACGCGGGCTTCTTTGGCGAGCTGTGCTCGA | 3971 |

| | | | |
|----|------|--|------|
| Db | 1750 | CGAGGACGCGGGCTATCGTGGCTTGCCACGACGGGGTTCCTTTGCGCAGCTGTGCTCGA | 1691 |
| Qy | 3972 | CGTTGTCACTGAAGCGGGAAGGAGCTGGCTGCTATTGGGGGAAGTGGCGGGCAGGATCT | 4031 |
| Db | 1690 | CGTTGTCACTGAAGCGGGAAGGAGCTGGCTGCTATTGGGGGAAGTGGCGGGCAGGATCT | 1631 |
| Qy | 4032 | CTGTGATCTCACCTTGTCTCCGAGAAAGTATCCATCATGCTGATGCAATGCGGCG | 4091 |
| Db | 1630 | CTGTGATCTCACCTTGTCTCCGAGAAAGTATCCATCATGCTGATGCAATGCGGCG | 1571 |
| Qy | 4092 | GCTGCATACGCTTGTATCCGGCTACCTGCGCCATTTCGACCAAGGGAACATCGCATCGA | 4151 |
| Db | 1570 | GCTGCATACGCTTGTATCCGGCTACCTGCGCCATTTCGACCAAGGGAACATCGCATCGA | 1511 |
| Qy | 4152 | GCAGACAGTACTCGGATGGAAGCCGGTCTTTGTCGATCAGGATGATCTGACGAGAGCA | 4211 |
| Db | 1510 | GCAGACAGTACTCGGATGGAAGCCGGTCTTTGTCGATCAGGATGATCTGACGAGAGCA | 1451 |
| Qy | 4212 | TCAGGGGCTCGCCAGCCGAACTGTCGCGAGCTCAAGCGCGCATGCCGACGGCGA | 4271 |
| Db | 1450 | TCAGGGGCTCGCCAGCCGAACTGTCGCGAGCTCAAGCGCGCATGCCGACGGCGA | 1391 |
| Qy | 4272 | GGATCTGCTGACCCATGCGGATGCTGCTTGCAGAAATATCATGCTGGAATGCGCG | 4331 |
| Db | 1390 | GGATCTGCTGACCCATGCGGATGCTGCTTGCAGAAATATCATGCTGGAATGCGCG | 1331 |
| Qy | 4332 | CTTTTCTGGAATCATGACATGTCGCGCGCTGGGTGTGGCGGACCGCTATCAGGACATAGC | 4391 |
| Db | 1330 | CTTTTCTGGAATCATGACATGTCGCGCGCTGGGTGTGGCGGACCGCTATCAGGACATAGC | 1271 |
| Qy | 4392 | GTTGGCTACCCGCTGATATTCGTAAGAGCTTGGCGGGAATGGCTGACCGCTTCTCGT | 4451 |
| Db | 1270 | GTTGGCTACCCGCTGATATTCGTAAGAGCTTGGCGGGAATGGCTGACCGCTTCTCGT | 1211 |
| Qy | 4452 | GCTTTACGGTATCGCGCTCCCGATTGCGAGCGCATCGCTTCTATCGCCTTCTTGACGA | 4511 |
| Db | 1210 | GCTTTACGGTATCGCGCTCCCGATTGCGAGCGCATCGCTTCTATCGCCTTCTTGACGA | 1151 |
| Qy | 4512 | GTTCTTCTGAGCGGAGCTCTGGGGTTCGATA | 4542 |
| Db | 1150 | GTTCTTCTGAGCGGAGCTCTGGGGTTCGATA | 1120 |

RESULT 5

US-10-523-682-1
; Sequence 1, Application US/10523682
; Publication No. US20060014149A1
; GENERAL INFORMATION:
; APPLICANT: Schnorr, Kirk
; TITLE OF INVENTION: Methods for rolling circle amplification and signal trapping of
; TITLE OF INVENTION: Libraries
; FILE REFERENCE: 10292.204-US
; CURRENT APPLICATION NUMBER: US/10/523,682
; CURRENT FILING DATE: 2005-02-01
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1
; LENGTH: 2403
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Plasmid pMHas5
US-10-523-682-1

Query Match 22.5%; Score 1166.2; DB 7; Length 2403;
Best Local Similarity 99.3%; Pred. No. 6.8e-267;
Matches 1171; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

| | | | |
|----|------|--|------|
| Qy | 3370 | CAAGCTTCAAGTCCGCGCAAGCACTCAGGGCGCAAGGGCTGCTTAAGGAACCGGAACACG | 3429 |
| Db | 8 | CTAGCTTCAAGTCCGCGCAAGCACTCAGGGCGCAAGGGCTGCTTAAGGAACCGGAACACG | 67 |

| | | | |
|----|------|--|------|
| Qy | 3430 | TAGAAAGCCAGTCCGCGAGAAACGGTGTCTGACCCCGGATGAATGTACGTACTACGGCTATC | 3489 |
| Db | 68 | TAGAAAGCCAGTCCGCGAGAAACGGTGTCTGACCCCGGATGAATGTACGTACTACGGCTATC | 127 |
| Qy | 3490 | TGGACAAGGGAAACCGAACGCGCAAGAGAAAGCAGGTACGTCCTTGCAGTGGGCTTACATGG | 3549 |
| Db | 128 | TGGACAAGGGAAACCGAACGCGCAAGAGAAAGCAGGTACGTCCTTGCAGTGGGCTTACATGG | 187 |
| Qy | 3550 | CGATAGCTAGACTGGGCGGTTTTATGCAAGCAGCAAGCGAACTGGAAATTTGCCAGCTGGGGCG | 3609 |
| Db | 188 | CGATAGCTAGACTGGGCGGTTTTATGCAAGCAGCAAGCGAACTGGAAATTTGCCAGCTGGGGCG | 247 |
| Qy | 3610 | CCCTCTGGTAAGGTTGGGAAGCCCTCGAAAGTAAACTGGAATGGCTTTCTTGGCGCAAGG | 3669 |
| Db | 248 | CCCTCTGGTAAGGTTGGGAAGCCCTCGAAAGTAAACTGGAATGGCTTTCTTGGCGCAAGG | 307 |
| Qy | 3670 | ATCTGATGGCGCAGGGGATCAAGATCTGATCAAGAGCAGGATGAGATCGTTTCCGATG | 3729 |
| Db | 308 | ATCTGATGGCGCAGGGGATCAAGATCTGATCAAGAGCAGGATGAGATCGTTTCCGATG | 367 |
| Qy | 3730 | ATTGAAACAAGATGATTCACGCGAGGTTCTCCGCGCGCTTGGGTGGAGAGCTATTTCGGC | 3789 |
| Db | 368 | ATTGAAACAAGATGATTCACGCGAGGTTCTCCGCGCGCTTGGGTGGAGAGCTATTTCGGC | 427 |
| Qy | 3790 | TATGACTGGGCACAAACAGACAATCGGCTGCTGTGATGCCCGCGTTTCCGGCTGTCAAGC | 3849 |
| Db | 428 | TATGACTGGGCACAAACAGACAATCGGCTGCTGTGATGCCCGCGTTTCCGGCTGTCAAGC | 487 |
| Qy | 3850 | CAGGGGCGCCGGTTCTTTTGTCAAGACCGACCTGTCCGGTCGCTGATGAATGAATGCAG | 3909 |
| Db | 488 | CAGGGGCGCCGGTTCTTTTGTCAAGACCGACCTGTCCGGTCGCTGATGAATGAATGCAG | 547 |
| Qy | 3910 | GACGAGCAGCGCGGCTATCGTGGCTGGCCACGACGCGGCGTTCTTTCGCGCAGCTGTGCTC | 3969 |
| Db | 548 | GACGAGCAGCGCGGCTATCGTGGCTGGCCACGACGCGGCGTTCTTTCGCGCAGCTGTGCTC | 607 |
| Qy | 3970 | GAGCTTGTCACTGAAGCGGAAAGGACTGTGCTGTCTATTGGGCGAAGTCCCGGGCAGGAT | 4029 |
| Db | 608 | GAGCTTGTCACTGAAGCGGAAAGGACTGTGCTGTCTATTGGGCGAAGTCCCGGGCAGGAT | 667 |
| Qy | 4030 | CTCCTGTCACTCACTTGTCTCTGCGAGAAAGTATCCATCATGGCTGATGCAATGCCG | 4089 |
| Db | 668 | CTCCTGTCACTCACTTGTCTCTGCGAGAAAGTATCCATCATGGCTGATGCAATGCCG | 727 |
| Qy | 4090 | CGGCTGATACGCTTGATCCGGCTACTCTGCCCATTCGACCACCAAGCGAAACATCGCATC | 4149 |
| Db | 728 | CGGCTGATACGCTTGATCCGGCTACTCTGCCCATTCGACCACCAAGCGAAACATCGCATC | 787 |
| Qy | 4150 | GAGCGACACGTACTCGGATGGAAAGCGGCTTTGTGATCAGGATGATCTGGACGAGAG | 4209 |
| Db | 788 | GAGCGACACGTACTCGGATGGAAAGCGGCTTTGTGATCAGGATGATCTGGACGAGAG | 847 |
| Qy | 4210 | CATCAGGGGCTCGGCGCAGCGGAACCTGTTGCCAGGCTCAAGCGCGGCATGCCCGACGGC | 4269 |
| Db | 848 | CATCAGGGGCTCGGCGCAGCGGAACCTGTTGCCAGGCTCAAGCGCGGCATGCCCGACGGC | 907 |
| Qy | 4270 | GAGGATCTCGTGTGACCCATGGCGATGCTGCTTGCAGAAATCATGGTGGAAATGGC | 4329 |
| Db | 908 | GAGGATCTCGTGTGACCCATGGCGATGCTGCTTGCAGAAATCATGGTGGAAATGGC | 967 |
| Qy | 4330 | CGCTTTTCTGGATTCATCGACTGTGGCGGCTGGGTGTGGCGGACCGCTATCAGGACATA | 4389 |
| Db | 968 | CGCTTTTCTGGATTCATCGACTGTGGCGGCTGGGTGTGGCGGACCGCTATCAGGACATA | 1027 |
| Qy | 4390 | CGGTTGGCTACCGGTGATATTTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGGCTCCTC | 4449 |
| Db | 1028 | CGGTTGGCTACCGGTGATATTTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGGCTCCTC | 1087 |
| Qy | 4450 | GTGCTTTACGGTATCGCGGCTCCCGATTTCCAGCGCATCGCCTTCTATCGCCTTCTTGAC | 4509 |
| Db | 1088 | GTGCTTTACGGTATCGCGGCTCCCGATTTCCAGCGCATCGCCTTCTATCGCCTTCTTGAC | 1147 |
| Qy | 4510 | GAGTTCTTCTGAGCGGACTCTGGGTTTCGATAAAATAA | 4548 |

```

Db      1148 GAGTTCCTTGACGGCGGACTCTGGGGTTCGCATGTAA 1186
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
RESULT 6
US-11-082-154A-31/c
; Sequence 31, Application US/11082154A
; Publication NO. US20060024820A1
; GENERAL INFORMATION:
; APPLICANT: Perkins, Edward
; APPLICANT: Perez, Carl
; APPLICANT: Lindenbaum, Michael
; APPLICANT: Greene, Amy
; APPLICANT: Leung, Josephine
; APPLICANT: Fleming, Elena
; APPLICANT: Stewart, Sandra
; APPLICANT: Shellard, Joan
; TITLE OF INVENTION: CHROMOSOME-BASED PLATFORMS
; FILE REFERENCE: 17084-022003 (420C)
; CURRENT APPLICATION NUMBER: US/11/082,154A
; CURRENT FILING DATE: 2005-03-15
; PRIOR APPLICATION NUMBER: 60/294,758
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: 60/366,891
; PRIOR FILING DATE: 2002-03-21
; PRIOR APPLICATION NUMBER: US 10/161,403
; PRIOR FILING DATE: 2002-05-30
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSeq For Windows Version 4.0
; SEQ ID NO 31
; LENGTH: 8136
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: pHE15 cosmid vector
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: GenBank X65279
; DATABASE ENTRY DATE: 1995-04-14
US-11-082-154A-31

```

| Query Match | 22.3% | Score 1154.4 | DB 12 | Length 8136 | |
|-----------------------|--------------|--|--------------|-------------|--------|
| Best Local Similarity | 99.8% | Pred. No. 8e-264 | | | |
| Matches 1166 | Conservative | 0 | Mismatches 1 | Indels 1 | Gaps 1 |
| Qy | 3375 | TTCAAGCTGCGCGAAGCAGCTCAGGGCGCAAGGCTGCTAAAGGAAGCGGAACACGCTAGAA | 3434 | | |
| Db | 4580 | TTCAAGCTGCGCGAAGCAGCTCAGGGCGCAAGGCTGCTAAAGGAAGCGGAACACGCTAGAA | 4521 | | |
| Qy | 3435 | AGCCAGTCCGCAAGAAACGGTGTGACCCCGATGAAATGTCAAGTTACTGGGCTATCTGGAC | 3494 | | |
| Db | 4520 | AGCCAGTCCGCAAGAAACGGTGTGACCCCGATGAAATGTCAAGTTACTGGGCTATCTGGAC | 4461 | | |
| Qy | 3495 | AAGGGAAACGCAAGCGCAAGGAAGCAAGGCTAGCTTCAGTGGGCTTACATGGCGATA | 3554 | | |
| Db | 4460 | AAGGGAAACGCAAGCGCAAGGAAGCAAGGCTAGCTTCAGTGGGCTTACATGGCGATA | 4401 | | |
| Qy | 3555 | GCTAGACTGGCGGTTTTTATGGACAGCAAGCGAATTCACAGCTGGGCGCGCCCTC | 3614 | | |
| Db | 4400 | GCTAGACTGGCGGTTTTTATGGACAGCAAGCGAATTCACAGCTGGGCGCGCCCTC | 4341 | | |
| Qy | 3615 | TGCTAAGTTTGGGAAGCCCTGCAAGCTTAACTGGATGGCTTCTTCGCGCCCAAGGATCTG | 3674 | | |
| Db | 4340 | TGCTAAGTTTGGGAAGCCCTGCAAGCTTAACTGGATGGCTTCTTCGCGCCCAAGGATCTG | 4281 | | |
| Qy | 3675 | ATCGCGCAGGGGATCAAGATCTGATCAAGAGA CAGGATCAGGATCGTTTCGCA TGA TTGA | 3734 | | |
| Db | 4280 | ATCGCGCAGGGGATCAAGATCTGATCAAGAGA CAGGATCAGGATCGTTTCGCA TGA TTGA | 4221 | | |
| Qy | 3735 | ACAAGATGATTCGACGCAGGTTCTCCGCGCTTCTGGTGGAGAGGCTATTTCGGCTATGA | 3794 | | |
| Db | 4220 | ACAAGATGATTCGACGCAGGTTCTCCGCGCTTCTGGTGGAGAGGCTATTTCGGCTATGA | 4161 | | |
| Qy | 3795 | CTGGGCACACAGACAATCGGCTGCTGATGTCGCGCGCTGTGTTCCGCGCTGTTCACGCGCAGGG | 3854 | | |

```
Db 4160 CTGGGCACACAGACAACTCGCTGCTGATGCGCGCGTGTTCGGCTGTTCAGCGCAGGG 4101
Qy 3855 GCGCCCGGTCTTTTGTCAAGACGACCTGTCCGGTCCCTGATGAATGCAATGCAAGGACGA 3914
Db 4100 GCGCCCGGTCTTTTGTCAAGACGACCTGTCCGGTCCCTGATGAATGCAATGCAAGGACGA 4041
Qy 3915 GGCAGCGGGTATCGTGGCTGGCCACACGCGGGGTCTTCGCGCAGCTGTGCTCGACGT 3974
Db 4040 GGCAGCGGGTATCGTGGCTGGCCACACGCGGGGTCTTCGCGCAGCTGTGCTCGACGT 3981
Qy 3975 TGTCACTGAAGCGGAAGGACTGTGCTCTATTTGGGCGAAGTGGCGGGCAGGATCTCCT 4034
Db 3980 TGTCACTGAAGCGGAAGGACTGTGCTCTATTTGGGCGAAGTGGCGGGCAGGATCTCCT 3921
Qy 4035 GTCACTCACTTGTCTCTGCGGAGAAAGTATCATATGCTGATGCTGATGCAATGCGCGGGCT 4094
Db 3920 GTCACTCACTTGTCTCTGCGGAGAAAGTATCATATGCTGATGCTGATGCAATGCGCGGGCT 3861
Qy 4095 GCATACGCTTGATCCGGCTACTCGCCCATTCGACCACCAAGCGAAACATCGATCGAGCG 4154
Db 3860 GCATACGCTTGATCCGGCTACTCGCCCATTCGACCACCAAGCGAAACATCGATCGAGCG 3801
Qy 4155 AGCAGCTACTCGGATGGAAGCGGTCTTGTTCGATCAGGATGATCTGGACGAAAGCATCA 4214
Db 3800 AGCAGCTACTCGGATGGAAGCGGTCTTGTTCGATCAGGATGATCTGGACGAAAGCATCA 3741
Qy 4215 GGGGCTCGGCCAGCGAACTGTTCCGACAGCTCAAGCGCGCAT-CCCGACGCGGAGGA 4274
Db 3740 GGGGCTCGGCCAGCGAACTGTTCCGACAGCTCAAGCGCGCAT-CCCGACGCGGAGGA 3682
Qy 4275 TCTGCTGTGACCCATGCGATGCTGCTGCGGAATATCATGTGGAAATGCGCGCTT 4334
Db 3681 TCTGCTGTGACCCATGCGATGCTGCTGCGGAATATCATGTGGAAATGCGCGCTT 3622
Qy 4335 TTCTGGATTTCATGACTGTGGCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTT 4394
Db 3621 TTCTGGATTTCATGACTGTGGCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTT 3562
Qy 4395 GCTACCCGCTGATATGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTCTCTCGTCT 4454
Db 3561 GCTACCCGCTGATATGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTCTCTCGTCT 3502
Qy 4455 TTACGGTATCGCGCTCCCGATTCGACGCGATCGCCTTCTATCGCCTTCTTGACGAGTT 4514
Db 3501 TTACGGTATCGCGCTCCCGATTCGACGCGATCGCCTTCTATCGCCTTCTTGACGAGTT 3442
Qy 4515 CTCTGAGCGGACTCTCGGGTTCGATA 4542
Db 3441 CTCTGAGCGGACTCTCGGGTTCGAAA 3414
```

```
RESULT 7
US-11-082-154A-118/c
; Sequence 118, Application US/11082154A
; Publication No. US20060024820A1
; GENERAL INFORMATION:
; APPLICANT: Perkins, Edward
; APPLICANT: Perez, Carl
; APPLICANT: Lindenbaum, Michael
; APPLICANT: Greene, Amy
; APPLICANT: Leung, Josephine
; APPLICANT: Fleming, Elena
; APPLICANT: Stewart, Sandra
; APPLICANT: Shellard, Joan
; TITLE OF INVENTION: CHROMOSOME-BASED PLATFORMS
; FILE REFERENCE: 17084-022003 (420C)
; CURRENT APPLICATION NUMBER: US/11/082,154A
; CURRENT FILING DATE: 2005-03-15
; PRIOR APPLICATION NUMBER: 60/294,758
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: 60/366,891
; PRIOR FILING DATE: 2002-03-21
```

```
; PRIOR APPLICATION NUMBER: US 10/161,403
; PRIOR FILING DATE: 2002-05-30
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 118
; LENGTH: 17384
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: pFK161 Plasmid
US-11-082-154A-118

Query Match      22.3%; Score 1154.4; DB 12; Length 17384;
Best Local Similarity 99.8%; Pred. No. 1.2e-263;
Matches 1166; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 3375 TTCAAGCTGCGCAGCACTCAGGGCGCAAGGGCTGTCTAAAGGAAGCGGAACACGTAGAA 3434
Db 2845 TTCAAGCTGCGCAGCACTCAGGGCGCAAGGGCTGTCTAAAGGAAGCGGAACACGTAGAA 2786
Qy 3435 AGCCAGTCCGACAGAAACGCTGCTGACCCCGATGAATGTCTGAGTCTGAGGCTATCTGGAC 3494
Db 2785 AGCCAGTCCGACAGAAACGCTGCTGACCCCGATGAATGTCTGAGTCTGAGGCTATCTGGAC 2726
Qy 3495 AAGGGAACCGCAAGCGCAAGAGAGAAAGCAGGTAGCTTGCAGTGGGCTTACATGGCGATA 3554
Db 2725 AAGGGAACCGCAAGCGCAAGAGAGAAAGCAGGTAGCTTGCAGTGGGCTTACATGGCGATA 2666
Qy 3555 GCTAGACTGGGCGGTTTTATGGACAGCAAGCGAAACCGGAATTGCCAGCTGGGGCGCCCTC 3614
Db 2665 GCTAGACTGGGCGGTTTTATGGACAGCAAGCGAAACCGGAATTGCCAGCTGGGGCGCCCTC 2606
Qy 3615 TGGTAAGTTGGGAAGCCCTGCAAGATGAATCTGATGCTCTTCTTCCGCCCAAGGATCTG 3674
Db 2605 TGGTAAGTTGGGAAGCCCTGCAAGATGAATCTGATGCTCTTCTTCCGCCCAAGGATCTG 2546
Qy 3675 ATGGCGCAGGGGATCAAGATCTGATCAAGAGACAGGATGAGGATCGTTTTCGATGATGA 3734
Db 2545 ATGGCGCAGGGGATCAAGATCTGATCAAGAGACAGGATGAGGATCGTTTTCGATGATGA 2486
Qy 3735 ACAAGATGGATTCGACGCAAGTTCTCCGGCGCTTGGGTGGAGAGGCTATTCGGCTATGA 3794
Db 2485 ACAAGATGGATTCGACGCAAGTTCTCCGGCGCTTGGGTGGAGAGGCTATTCGGCTATGA 2426
Qy 3795 CTGGGCAACAGACATCGCTGCTGATGCGCGCTGTTCCGGCTGTCAGCGCAGGG 3854
Db 2425 CTGGGCAACAGACATCGCTGCTGATGCGCGCTGTTCCGGCTGTCAGCGCAGGG 2366
Qy 3855 GCGCCCGGTTCTTTTGTCAAGACCGGACCTGTCCGGTGCCTGATGAATGAACTGCAGGACGA 3914
Db 2365 GCGCCCGGTTCTTTTGTCAAGACCGGACCTGTCCGGTGCCTGATGAATGAACTGCAGGACGA 2306
Qy 3915 GGCAGCGCGCTATCGTGGCTGGCCACGACGGGCGTTCTTTCGCGAGCTGTGCTCGACGT 3974
Db 2305 GGCAGCGCGCTATCGTGGCTGGCCACGACGGGCGTTCTTTCGCGAGCTGTGCTCGACGT 2246
Qy 3975 TGTCACTGAAGCGGAAGGACTGGCTGCTATTTGGGCGAAGTGGCGGGCAGGATCTCCT 4034
Db 2245 TGTCACTGAAGCGGAAGGACTGGCTGCTATTTGGGCGAAGTGGCGGGCAGGATCTCCT 2186
Qy 4035 GTCACTCACCTTGTCTCTCCGAGAAAGTATCCATCATGTGCTGATGCAATGCGCGGCT 4094
Db 2185 GTCACTCACCTTGTCTCTCCGAGAAAGTATCCATCATGTGCTGATGCAATGCGCGGCT 2126
Qy 4095 GCATACGCTTGATCCGGCTACCTGCCCATTCGACCAACAGCGGAACATCCATCGAGCG 4154
Db 2125 GCATACGCTTGATCCGGCTACCTGCCCATTCGACCAACAGCGGAACATCCATCGAGCG 2066
Qy 4155 AGCAGGCTACTCGGATGGAAGCGGCTGTGCTGATCAGGATGATCTGGAAGAGGATCA 4214
Db 2065 AGCAGGCTACTCGGATGGAAGCGGCTGTGCTGATCAGGATGATCTGGAAGAGGATCA 2006
Qy 4215 GGGGCTCGGCCAGCGAACTGTTTCGCGCAGGCTCAAGGCGCGCATGCTCCCGACGCGGAGGA 4274
```

Db 2005 GGGGCTCGCGCCAGCCGAACTGTCGCGAGGCTCAAGGCGGCGCAT-CCCGACGGCGAGGA 1947
Qy 4275 TCTCGTGTGACCCATGCGGATGCTGCTTGCAGAAATATCATGTGGGAAATGCGCGCTT 4334
Db 1946 TCTCGTGTGACCCATGCGGATGCTGCTTGCAGAAATATCATGTGGGAAATGCGCGCTT 1887
Qy 4335 TCTCGATTCATGACTGTGCGCGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTT 4394
Db 1886 TTCTGGATTTCATGACTGTGCGCGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTT 1827
Qy 4395 GCCTACCCGTCATATGCTGAAGAGCTTGGGGGGAATGGGCTGACCGCTTCTCGTGTCT 4454
Db 1826 GCCTACCCGTCATATGCTGAAGAGCTTGGGGGGAATGGGCTGACCGCTTCTCGTGTCT 1767
Qy 4455 TTACGGTATCGCGCTCCCGATTCGACGCGATCGCCCTTCTATCGCCCTTCTTGACGAGTT 4514
Db 1766 TTACGGTATCGCGCTCCCGATTCGACGCGATCGCCCTTCTATCGCCCTTCTTGACGAGTT 1707
Qy 4515 CTTCTGAGCGGAGCTCTGGGGTTTGATA 4542
Db 1706 CTTCTGAGCGGAGCTCTGGGGTTTGATA 1679

RESULT 8

US-10-525-710-66
; Sequence 66, Application US/10525710
; Publication No. US20050260721A1
; GENERAL INFORMATION:
; APPLICANT: Kroger, Burkhard
; APPLICANT: Zeider, Oskar
; APPLICANT: Kolpprogge, Corinna
; APPLICANT: Schroder, Hartwig
; APPLICANT: Hafner, Stefan
; TITLE OF INVENTION: Method for Zymotic Production of Fine Chemicals Containing
; TITLE OF INVENTION: Sulphur (mety)
; FILE REFERENCE: 13111-00006-US
; CURRENT APPLICATION NUMBER: US/10/525,710
; CURRENT FILING DATE: 2005-02-24
; PRIOR APPLICATION NUMBER: PCT/EP 2003/009453
; PRIOR FILING DATE: 2003-08-25
; PRIOR APPLICATION NUMBER: DE 102 39 082.7
; PRIOR FILING DATE: 2002-08-26
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 66
; LENGTH: 4323
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Description of the artificial sequence: plasmid
US-10-525-710-66

Query Match 22.0%; Score 1136.4; DB 8; Length 4323;
Best Local Similarity 99.9%; Pred. No. 1.1e-259;
Matches 1137; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 3405 GGGCTGCTAAAGGAAGCGGAACAGCTAGAAAGCCAGTCCGAGAAACCGTCTGACCCCG 3464
Db 1359 GGGCTGCTAAAGGAAGCGGAACAGCTAGAAAGCCAGTCCGAGAAACCGTCTGACCCCG 1418
Qy 3465 GATGAATGTACGTACTGCGGTATCTGGAAGGGAACCAAGCGCAAGAAAGCA 3524
Db 1419 GATGAATGTACGTACTGCGGTATCTGGAAGGGAACCAAGCGCAAGAAAGCA 1478
Qy 3525 GGTAGCTTGCATGGGCTTACATGCGGATAGTACATGCGGCGTTTATGGAAGCAAG 3584
Db 1479 GGTAGCTTGCATGGGCTTACATGCGGATAGTACATGCGGCGTTTATGGAAGCAAG 1538
Qy 3585 CGAACCGGAATTGCCAGCTGGGCGCCCTCTGTTAAGTTGGAAGCCCTGCAAGTAA 3644
Db 1539 CGAACCGGAATTGCCAGCTGGGCGCCCTCTGTTAAGTTGGAAGCCCTGCAAGTAA 1598

RESULT 9

US-10-525-674-58
; Sequence 58, Application US/10525674
; Publication No. US20060003425A1
; GENERAL INFORMATION:
; APPLICANT: Kroger, Burkhard
; APPLICANT: Zeider, Oskar
; APPLICANT: Kolpprogge, Corinna
; APPLICANT: Schroder, Hartwig
; APPLICANT: Hafner, Stefan
; TITLE OF INVENTION: Method for Zymotic Production of Fine Chemicals Containing
; TITLE OF INVENTION: Sulphur (Meta)

Qy 3645 CTGATGGCTTTCTTGCGCCCAAGGATCTGATGGCGAGGGGATCAAGATCTGATCAAGA 3704
Db 1599 CTGATGGCTTTCTTGCGCCCAAGGATCTGATGGCGAGGGGATCAAGATCTGATCAAGA 1658
Qy 3705 GACAGGATGAGGATCGTTTTCGATGATGAACAGATGATGACCGCAGGTTCTCGGC 3764
Db 1659 GACAGGATGAGGATCGTTTTCGATGATGAACAGATGATGACCGCAGGTTCTCGGC 1718
Qy 3765 CGCTTGGGTGAGAGGCTATTTCGCTATGACTGGGCAACAGACAAATCGGCTGCTGA 3824
Db 1719 CGCTTGGGTGAGAGGCTATTTCGCTATGACTGGGCAACAGACAAATCGGCTGCTGA 1778
Qy 3825 TGCCCGCTGTTCGCGCTGTGACGCGAGGGGCGCCCGTTCTTTTGTCAAGACGACCT 3884
Db 1779 TGCCCGCTGTTCGCGCTGTGACGCGAGGGGCGCCCGTTCTTTTGTCAAGACGACCT 1838
Qy 3885 GTCCGCTGCCCTGAATGAATGACGAGGAGCGCGCTATCTGCGCTGGCGGACGAC 3944
Db 1839 GTCCGCTGCCCTGAATGAATGACGAGGAGCGCGCTATCTGCGCTGGCGGACGAC 1898
Qy 3945 GGGCGTTCTTTGCGCAGCTGTCTCGAGCTTGTCTCACTGAAGCGGGAAGGACTTGGCTGCT 4004
Db 1899 GGGCGTTCTTTGCGCAGCTGTCTCGAGCTTGTCTCACTGAAGCGGGAAGGACTTGGCTGCT 1958
Qy 4005 ATTGGGCAAGTGCCGGGCGAGGATCTCTGCTCATCTCACTGCTCTGCGGAGAAAGT 4064
Db 1959 ATTGGGCAAGTGCCGGGCGAGGATCTCTGCTCATCTCACTGCTCTGCGGAGAAAGT 2018
Qy 4065 ATCCATCATGGCTGATGCAATGCGCGGCTGCATACGCTTGATCGGCTACCTGCCAAT 4124
Db 2019 ATCCATCATGGCTGATGCAATGCGCGGCTGCATACGCTTGATCGGCTACCTGCCAAT 2078
Qy 4125 CGACCAACGAGCAACATCGCATCGAGCGAGCACTACTCGGATGGAAGCGGCTTGTGT 4184
Db 2079 CGACCAACGAGCAACATCGCATCGAGCGAGCACTACTCGGATGGAAGCGGCTTGTGT 2138
Qy 4185 CGATCAGATGATCTGACGAGAGCATCAGGGGCTCGCGCAGCGAACTGTTCGCCAG 4244
Db 2139 CGATCAGATGATCTGACGAGAGCATCAGGGGCTCGCGCAGCGAACTGTTCGCCAG 2198
Qy 4245 GCTCAAGCGCGCATGCCGAGGAGTCTGCTGAGCCCATGCGGATGCTGCTT 4304
Db 2199 GCTCAAGCGCGCATGCCGAGGAGTCTGCTGAGCCCATGCGGATGCTGCTT 2258
Qy 4305 GCCGAATATCATGTGAAATGCGCGCTTTCTGGAATCATCGACTGTGCGCGCTGGG 4364
Db 2259 GCCGAATATCATGTGAAATGCGCGCTTTCTGGAATCATCGACTGTGCGCGCTGGG 2318
Qy 4365 TGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATGCTGAAGAGCTTGG 4424
Db 2319 TGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATGCTGAAGAGCTTGG 2378
Qy 4425 CGGCGAATGGCTGACCGCTTCTGCTGCTTACGCTTACGCTTCCGCTCCCGATTCGACGCG 4484
Db 2379 CGGCGAATGGCTGACCGCTTCTGCTGCTTACGCTTACGCTTCCGCTCCCGATTCGACGCG 2438
Qy 4485 CATCGCTTCTATCGCTTCTTGAAGAGTCTTCTGAGCGGAGCTCTGGGGTTCGATA 4542
Db 2439 CATCGCTTCTATCGCTTCTTGAAGAGTCTTCTGAGCGGAGCTCTGGGGTTCGATA 2496

```
; FILE REFERENCE: 13111-00002-US
; CURRENT APPLICATION NUMBER: US/10/525,674
; CURRENT FILING DATE: 2005-02-24
; PRIOR APPLICATION NUMBER: PCT/EP 2003/009452
; PRIOR FILING DATE: 2003-08-26
; PRIOR APPLICATION NUMBER: DE 102 39 073.8
; PRIOR FILING DATE: 2002-08-26
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 58
; LENGTH: 4323
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Description of the artificial sequence: plasmid
US-10-525-674-58

Query Match      22.0%; Score 1136.4; DB 8; Length 4323;
Best Local Similarity 99.9%; Pred. No. 1.1e-259;
Matches 1137; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3405 GGGCTGCTAAAGGAAGCGGAACACGTAGAAAGCCAGTCCGAGAAACCGGTGCTGACCCCG 3464
Db 1359 GGGCTGCTAAAGGAAGCGGAACACGTAGAAAGCCAGTCCGAGAAACCGGTGCTGACCCCG 1418

Qy 3465 GATGAATGTCAGTACTGGGCTATCTGCACAGGGAACCAAGCGCAAGAGAAAGCA 3524
Db 1419 GATGAATGTCAGTACTGGGCTATCTGCACAGGGAACCAAGCGCAAGAGAAAGCA 1478

Qy 3525 GGTAGCTTGCACTGGGCTTACATGCGGATAGTACGCTGGGGTATTTATGCACAGCAAG 3584
Db 1479 GGTAGCTTGCACTGGGCTTACATGCGGATAGTACGCTGGGGTATTTATGCACAGCAAG 1538

Qy 3585 CGAACCGGAATTGCCAGCTGGGGGCCCTCTGTTAAGTTGGGAAGCCCTCAAGTAA 3644
Db 1539 CGAACCGGAATTGCCAGCTGGGGGCCCTCTGTTAAGTTGGGAAGCCCTCAAGTAA 1598

Qy 3645 CTGGATGCTTTCTTGGCCCAAGATCTGATGGCGAGGGGATCAAGATCTGATCAAGA 3704
Db 1599 CTGGATGCTTTCTTGGCCCAAGATCTGATGGCGAGGGGATCAAGATCTGATCAAGA 1658

Qy 3705 GACAGGATGAGGATCGTTTCCGATGATTGAACAAGATGATTGACGAGGTTCTCCGGC 3764
Db 1659 GACAGGATGAGGATCGTTTCCGATGATTGAACAAGATGATTGACGAGGTTCTCCGGC 1718

Qy 3765 CGCTTGGGTGAGAGGCTATTGCGCTATGACTGGGCAACAAGCAATCGGCTGCTCTGA 3824
Db 1719 CGCTTGGGTGAGAGGCTATTGCGCTATGACTGGGCAACAAGCAATCGGCTGCTCTGA 1778

Qy 3825 TGCCGCGCTGTTCCGGCTGTCAGCGAGGGCGCCGGTCTTTTGTGTCAGACCGACCT 3884
Db 1779 TGCCGCGCTGTTCCGGCTGTCAGCGAGGGCGCCGGTCTTTTGTGTCAGACCGACCT 1838

Qy 3885 GTCCGGTGCCTGAATGAATTCAGAGCAGCGCGGCTATCGTGGCTGGCCACGAC 3944
Db 1839 GTCCGGTGCCTGAATGAATTCAGAGCAGCGCGGCTATCGTGGCTGGCCACGAC 1898

Qy 3945 GGGCGGTTCTTTCGCGAGCTGTGCTCGACGTTGTCTA CTGAAGCGGGAAGGGAATGCTGCT 4004
Db 1899 GGGCGGTTCTTTCGCGAGCTGTGCTCGACGTTGTCTA CTGAAGCGGGAAGGGAATGCTGCT 1958

Qy 4005 ATTGGGCGAAGTCCGGGGCAGGATCTCTGTCTCATCTCACTTGTCTCTCGGAGAAAGT 4064
Db 1959 ATTGGGCGAAGTCCGGGGCAGGATCTCTGTCTCATCTCACTTGTCTCTCGGAGAAAGT 2018

Qy 4065 ATCCATCATGGCTGATGCAATGCGCGGCTGTCATACGCTTGATCGGGCTACTGCGCCATT 4124
Db 2019 ATCCATCATGGCTGATGCAATGCGCGGCTGTCATACGCTTGATCGGGCTACTGCGCCATT 2078

Qy 4125 CGACCAACCAAGCGAAACATCGCATCGAGCGAGCACGTACTCGATGGAAGCGGCTTTGT 4184
Db 2079 CGACCAACCAAGCGAAACATCGCATCGAGCGAGCACGTACTCGATGGAAGCGGCTTTGT 2138

; FILE REFERENCE: 13111-00006-US
; Sequence 65, Application US/10525710
; Publication No. US20050260721A1
; GENERAL INFORMATION:
; APPLICANT: Kroger, Burkhard
; APPLICANT: Zelder, Oskar
; APPLICANT: Kolpprogge, Corinna
; APPLICANT: Schroder, Hartwig
; APPLICANT: Hafner, Stefan
; TITLE OF INVENTION: Method for Zymotic Production of Fine Chemicals Containing
; TITLE OF INVENTION: Sulphur (mety)
; FILE REFERENCE: 13111-00006-US
; CURRENT APPLICATION NUMBER: US/10/525,710
; CURRENT FILING DATE: 2005-02-24
; PRIOR APPLICATION NUMBER: PCT/EP 2003/009453
; PRIOR FILING DATE: 2003-08-26
; PRIOR APPLICATION NUMBER: DE 102 39 082.7
; PRIOR FILING DATE: 2002-08-26
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 65
; LENGTH: 5091
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Description of the artificial sequence: plasmid
US-10-525-710-65

Query Match      22.0%; Score 1136.4; DB 8; Length 5091;
Best Local Similarity 99.9%; Pred. No. 1.2e-259;
Matches 1137; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3405 GGGCTGCTAAAGGAAGCGGAACACGTAGAAAGCCAGTCCGAGAAACCGGTGCTGACCCCG 3464
Db 358 GGGCTGCTAAAGGAAGCGGAACACGTAGAAAGCCAGTCCGAGAAACCGGTGCTGACCCCG 417

Qy 3465 GATGAATGTCAGTACTGGGCTATCTGCACAGGGAACCAAGCGCAAGAGAAAGCA 3524
Db 418 GATGAATGTCAGTACTGGGCTATCTGCACAGGGAACCAAGCGCAAGAGAAAGCA 477

Qy 3525 GGTAGCTTGCACTGGGCTTACATGCGGATAGTACGCTGGGGTATTTATGCACAGCAAG 3584
Db 478 GGTAGCTTGCACTGGGCTTACATGCGGATAGTACGCTGGGGTATTTATGCACAGCAAG 537

Qy 3585 CGAACCGGAATTGCCAGCTGGGGGCCCTCTGTTAAGTTGGGAAGCCCTCAAGTAA 3644
Db 538 CGAACCGGAATTGCCAGCTGGGGGCCCTCTGTTAAGTTGGGAAGCCCTCAAGTAA 597
```


QY 4185 CGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCGCGCAGCCGAACTGTTGCGCAG 4244
DB 1138 CGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCGCGCAGCCGAACTGTTGCGCAG 1197
QY 4245 GCTCAAGCGCGCATGCCCGACGGCGAGGATCTCGTGTGACCCCATGCGCATGCTGCTT 4304
DB 1198 GCTCAAGCGCGCATGCCCGACGGCGAGGATCTCGTGTGACCCCATGCGCATGCTGCTT 1257
QY 4305 GCCGAATATCATGCTGGAAATGGCCGCTTTCTGGATTCATCGACTGTGCGCGGCTGGG 4364
DB 1258 GCCGAATATCATGCTGGAAATGGCCGCTTTCTGGATTCATCGACTGTGCGCGGCTGGG 1317
QY 4365 TGTGCGGACCGCTATCAGGACATAGCGTTCGCTACCCGTGATATGCTGAAGAGCTTGG 4424
DB 1318 TGTGCGGACCGCTATCAGGACATAGCGTTCGCTACCCGTGATATGCTGAAGAGCTTGG 1377
QY 4425 CGCGGAATGGGCTGACCGCTTCTCGTGTCTTTACGGTATCGCCGCTCCCGATTCGACGC 4484
DB 1378 CGCGGAATGGGCTGACCGCTTCTCGTGTCTTTACGGTATCGCCGCTCCCGATTCGACGC 1437
QY 4485 CATGCCCTTCTATCGCCCTTCTGACGAGTTCCTTGACGGGAGCTCTGGGGTTTGATA 4542
DB 1438 CATGCCCTTCTATCGCCCTTCTGACGAGTTCCTTGACGGGAGCTCTGGGGTTTGATA 1495

RESULT 12
US-10-525-710-69
; Sequence 69, Application US/10525710
; Publication No. US20050260721A1
; GENERAL INFORMATION:
; APPLICANT: Kroger, Burkhard
; APPLICANT: Zelder, Oskar
; APPLICANT: Kolpprogge, Corinna
; APPLICANT: Schroder, Hartwig
; TITLE OF INVENTION: Method for Zymotic Production of Fine Chemicals Containing
; TITLE OF INVENTION: Sulphur (metY)
; FILE REFERENCE: 13111-00006-US
; CURRENT APPLICATION NUMBER: US/10/525,710
; CURRENT FILING DATE: 2005-02-24
; PRIOR APPLICATION NUMBER: PCT/EP 2003/009453
; PRIOR FILING DATE: 2003-08-26
; PRIOR APPLICATION NUMBER: DE 102 39 082.7
; PRIOR FILING DATE: 2002-08-26
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 69
; LENGTH: 5860
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURES:
; OTHER INFORMATION: Description of the artificial sequence: plasmid
US-10-525-710-69

Query Match 22.0%; Score 1136.4; DB 8; Length 5860;
Best Local Similarity 99.9%; Pred. No. 1.3e-259;
Matches 1137; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3405 GGGCTGCTAAAGGAAGCGGAACACCTAGAAAAGCCAGTCCGAGAAACCGGTGCTGACCCCG 3464
DB 1652 GGGCTGCTAAAGGAAGCGGAACACCTAGAAAAGCCAGTCCGAGAAACCGGTGCTGACCCCG 1711
QY 3465 GATGAATGTAGCTACTGGGCTATCTGGACGAAGGAAACCGCAAGCAAGCAAGCAAGCA 3524
DB 1712 GATGAATGTAGCTACTGGGCTATCTGGACGAAGGAAACCGCAAGCAAGCAAGCAAGCA 1771
QY 3525 GGTAGCTTGCAGTGGGCTTACATGCGGATAGCTAGCTGGGCGGTTTTATGACAGCAAG 3584
DB 1772 GGTAGCTTGCAGTGGGCTTACATGCGGATAGCTAGCTGGGCGGTTTTATGACAGCAAG 1831
QY 3585 GCAACCGGAATTCAGCAGCTGGGGCGCCCTCTGTTAAGGTTGGGAAGCCCTGCAAGATAA 3644

DB 1832 CGAACCGGAATTCAGCTGGGGCGCCCTCTGTGTAAAGTTGGGAAGCCCTGCAAGATAA 1891
QY 3645 CTGGATGGCTTTCTTGGCCGCAAGGATCTGATGGCGCAGGGGATCAAGATCTGATCAAGA 3704
DB 1892 CTGGATGGCTTTCTTGGCCGCAAGGATCTGATGGCGCAGGGGATCAAGATCTGATCAAGA 1951
QY 3705 GACAGGATGAGGATCGTTTTCGCATGATTGAACAAGATGGATTCACGACAGGTTCTCGGC 3764
DB 1952 GACAGGATGAGGATCGTTTTCGCATGATTGAACAAGATGGATTCACGACAGGTTCTCGGC 2011
QY 3765 CGCTTGGGTGGAAGGCTATTCGGCTATGACTGGGCAACAAGCAATTCGGCTCTCTGA 3824
DB 2012 CGCTTGGGTGGAAGGCTATTCGGCTATGACTGGGCAACAAGCAATTCGGCTCTCTGA 2071
QY 3825 TGCGCCGCTGTTCCGGCTGTTCAGGCGAGGGCGCCCGGTTCTTTTCTCAAGACCGACCT 3884
DB 2072 TGCGCCGCTGTTCCGGCTGTTCAGGCGAGGGCGCCCGGTTCTTTTCTCAAGACCGACCT 2131
QY 3885 GTCCGGTGCCTCTGAATGAACCTGCAGGACGAGGACGCGGCTATCTGTGGCTGGCCACGAC 3944
DB 2132 GTCCGGTGCCTCTGAATGAACCTGCAGGACGAGGACGCGGCTATCTGTGGCTGGCCACGAC 2191
QY 3945 GGGGTTCTTGGCGAGCTGTCTCGACGTTGTCTCAAGCGGGAAGGAGTCTGGTGTCT 4004
DB 2192 GGGGTTCTTGGCGAGCTGTCTCGACGTTGTCTCAAGCGGGAAGGAGTCTGGTGTCT 2251
QY 4005 ATTGGGCGAAGTCCGGGGCAGGATCTCTGTCTCATCTCACCTTGTCTCTCGCCGAGAAAGT 4064
DB 2252 ATTGGGCGAAGTCCGGGGCAGGATCTCTGTCTCATCTCACCTTGTCTCTCGCCGAGAAAGT 2311
QY 4065 ATCCATCATGGCTGATGCAATGGCGGCTGCATACCGTTTGTATCCGGCTACCTGCCCAT 4124
DB 2312 ATCCATCATGGCTGATGCAATGGCGGCTGCATACCGTTTGTATCCGGCTACCTGCCCAT 2371
QY 4125 CGACCAAGCGAAACATCGCATTCGACGAGCAGCCTGCTCGGATGGAAGCGGCTGTGT 4184
DB 2372 CGACCAAGCGAAACATCGCATTCGACGAGCAGCCTGCTCGGATGGAAGCGGCTGTGT 2431
QY 4185 CGATCAGGATGATTCGACGAAGGATCAGGGGCTCGCGCCAGCCGAACTGTTTCGCCAG 4244
DB 2432 CGATCAGGATGATTCGACGAAGGATCAGGGGCTCGCGCCAGCCGAACTGTTTCGCCAG 2491
QY 4245 GCTCAAGCGCGATGCCCGACCGAGGATCTCGTGTGACCCATCGCGATGCTGCTT 4304
DB 2492 GCTCAAGCGCGATGCCCGACCGAGGATCTCGTGTGACCCATCGCGATGCTGCTT 2551
QY 4305 GCCGAATATCATGGTGAATAATGGCCGCTTTCTGGATTCATCGACTGTGGCGGCTGGG 4364
DB 2552 GCCGAATATCATGGTGAATAATGGCCGCTTTCTGGATTCATCGACTGTGGCGGCTGGG 2611
QY 4365 TGTGGCGAGCGCTATCAGGACATAGCGTTGGCTTACCGGTGATATTCGTAAGAGCTTGG 4424
DB 2612 TGTGGCGAGCGCTATCAGGACATAGCGTTGGCTTACCGGTGATATTCGTAAGAGCTTGG 2671
QY 4425 CGCGGAATGGGCTGACCGCTTCTCGTGTCTTACGGTATCGCCGCTCCGATTCGACGCG 4484
DB 2672 CGCGGAATGGGCTGACCGCTTCTCGTGTCTTACGGTATCGCCGCTCCGATTCGACGCG 2731
QY 4485 CATCGCTTCTATCGCCTTCTTGAAGGATTCCTGAGCGGAGCTCTGGGGTTTGATA 4542
DB 2732 CATCGCTTCTATCGCCTTCTTGAAGGATTCCTGAGCGGAGCTCTGGGGTTTGATA 2789

RESULT 13
US-10-525-710-74
; Sequence 74, Application US/10525710
; Publication No. US20050260721A1
; GENERAL INFORMATION:
; APPLICANT: Kroger, Burkhard
; APPLICANT: Zelder, Oskar
; APPLICANT: Kolpprogge, Corinna
; APPLICANT: Schroder, Hartwig
; APPLICANT: Hafner, Stefan

;; TITLE OF INVENTION: Method for Zymotic Production of Fine Chemicals Containing
;; FILE REFERENCE: Sulphur (mef)
;; CURRENT APPLICATION NUMBER: US/10/525,710
;; CURRENT FILING DATE: 2005-02-24
;; PRIOR APPLICATION NUMBER: PCT/EP 2003/009453
;; PRIOR FILING DATE: 2003-08-26
;; PRIOR APPLICATION NUMBER: DE 102 39 082.7
;; PRIOR FILING DATE: 2002-08-26
;; SOFTWARE: PatentIn version 3.3
;; SEQ ID NO 74
;; LENGTH: 5860
;; TYPE: DNA
;; ORGANISM: Artificial sequence
;; FEATURE:
;; OTHER INFORMATION: Description of the artificial sequence: plasmid
US-10-525-710-74

Query Match 22.0%; Score 1136.4; DB 8; Length 5860;
Best Local Similarity 99.9%; Pred. No. 1.3e-259;
Matches 1137; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3405 GGGCTGCTAAAGGAAAGCGGAAACACCTAGAAAGCCAGTCCGAGAAACGGTGTGACCCCG 3464
DB 1652 GGGCTGCTAAAGGAAAGCGGAAACACCTAGAAAGCCAGTCCGAGAAACGGTGTGACCCCG 1711

QY 3465 GATGAATGTACGCTACTGGCTTATCTGACAAAGGAAACGCAAGCGCAAGAAAGCA 3524
DB 1712 GATGAATGTACGCTACTGGCTTATCTGACAAAGGAAACGCAAGCGCAAGAAAGCA 1771

QY 3525 GGTAGCTTCAGTGGCTTACATGCGGATAGTACAGTGGCGGTTTTATGACAGCAAG 3584
DB 1772 GGTAGCTTCAGTGGCTTACATGCGGATAGTACAGTGGCGGTTTTATGACAGCAAG 1831

QY 3585 CGAACCGGAATGCCAGCTGGGCGCCCTCTGTAAGTTGGGAAGCCCTGCAAGTAAA 3644
DB 1832 CGAACCGGAATGCCAGCTGGGCGCCCTCTGTAAGTTGGGAAGCCCTGCAAGTAAA 1891

QY 3645 CTGGATGCTTTCTTGGCCCAAGATCTGATGGCGCAGGGGATCAAGATCTGATCAAGA 3704
DB 1892 CTGGATGCTTTCTTGGCCCAAGATCTGATGGCGCAGGGGATCAAGATCTGATCAAGA 1951

QY 3705 GACAGGATGAGGATCGTTTCGATGATGAACAGATGATGACGAGGTTCTCGGC 3764
DB 1952 GACAGGATGAGGATCGTTTCGATGATGAACAGATGATGACGAGGTTCTCGGC 2011

QY 3765 CGCTTGGGTGAGAGGCTATTGCGCTATGATCGGCAACAGCAATCGGCTGCTCTGA 3824
DB 2012 CGCTTGGGTGAGAGGCTATTGCGCTATGATCGGCAACAGCAATCGGCTGCTCTGA 2071

QY 3825 TGCCCGCTGTTCCGGCTGTGAGCGCAGGGGGCCGGTCTTTTGTGCAAGCCGACCT 3884
DB 2072 TGCCCGCTGTTCCGGCTGTGAGCGCAGGGGGCCGGTCTTTTGTGCAAGCCGACCT 2131

QY 3885 GTCCGGTCCCTGAATGAATGACGAGCAGCGGCTATCGTGGCTGGCCACGAC 3944
DB 2132 GTCCGGTCCCTGAATGAATGACGAGCAGCGGCTATCGTGGCTGGCCACGAC 2191

QY 3945 GGGCGTTCCTTGGCGAGCTGTGCTGAGTGTGCTACTGAAGCGGAAAGGACTGGCTGCT 4004
DB 2192 GGGCGTTCCTTGGCGAGCTGTGCTGAGTGTGCTACTGAAGCGGAAAGGACTGGCTGCT 2251

QY 4005 ATTGGGCGAAGTGGCGGCGAGGATCTCTGTCATCTCACTTGTCTGCTCGGAGAGT 4064
DB 2252 ATTGGGCGAAGTGGCGGCGAGGATCTCTGTCATCTCACTTGTCTGCTCGGAGAGT 2311

QY 4065 ATCCATCATGCTGATGCAATGCGCGCTGCTATACGCTTGTATCGGCTACTGCCCCATT 4124
DB 2312 ATCCATCATGCTGATGCAATGCGCGCTGCTATACGCTTGTATCGGCTACTGCCCCATT 2371

QY 4125 CGAACCAAGCGAAACATCGCATCGAGCGACGATGCTCGGATGGGAAGCCGCTTTGT 4184

DB 2372 CGAACCAAGCGAAACATCGCATCGAGCGAGCAGTACTCGGATGGAGCCGCTCTTGT 2431
QY 4185 CGATCAGGATGATCTGGAAGAGCATCAGGGGCTCGCGCAGCCGAACTGTTCGCCAG 4244
DB 2432 CGATCAGGATGATCTGGAAGAGCATCAGGGGCTCGCGCAGCCGAACTGTTCGCCAG 2491
QY 4245 GCTCAAGCGCGCATGCCCGAGCGGAGGATCTGTCGTGTGACCCATGCGGATGCTGCTT 4304
DB 2492 GCTCAAGCGCGCATGCCCGAGCGGAGGATCTGTCGTGTGACCCATGCGGATGCTGCTT 2551
QY 4305 GCCGAATATCATGTGGAAATGCGCTTTCTTGGGATTCATCGACTGTGCGCGCTGG 4364
DB 2552 GCCGAATATCATGTGGAAATGCGCTTTCTTGGGATTCATCGACTGTGCGCGCTGG 2611
QY 4365 TGTGCGGACCGCTATCAGGACATAGCGTTGGCTTACCGGTATGCTGAAGAGCTTGG 4424
DB 2612 TGTGCGGACCGCTATCAGGACATAGCGTTGGCTTACCGGTATGCTGAAGAGCTTGG 2671

QY 4425 CGGCGAATGGCTGACCGCTTCTCGTGTCTTACGGTATCGCGCTCCCGATTCGACGG 4484
DB 2672 CGGCGAATGGCTGACCGCTTCTCGTGTCTTACGGTATCGCGCTCCCGATTCGACGG 2731

QY 4485 CATCGCTTCTATCGCTTCTTGACGAGTCTTCTGAGCGGAGCTCTGCGGCTTCGATA 4542
DB 2732 CATCGCTTCTATCGCTTCTTGACGAGTCTTCTGAGCGGAGCTCTGCGGCTTCGATA 2789

RESULT 14
US-10-525-674-61
;; Sequence 61, Application US/10525674
;; Publication No. US20060003425A1
;; GENERAL INFORMATION:
;; APPLICANT: Kroger, Burkhard
;; APPLICANT: Zelder, Oskar
;; APPLICANT: Kolprogge, Corinna
;; APPLICANT: Schroder, Hartwig
;; APPLICANT: Hafner, Stefan
;; TITLE OF INVENTION: Method for Zymotic Production of Fine Chemicals Containing
;; FILE REFERENCE: Sulphur (Meta)
;; CURRENT APPLICATION NUMBER: US/10/525,674
;; CURRENT FILING DATE: 2005-02-24
;; PRIOR APPLICATION NUMBER: PCT/EP 2003/009452
;; PRIOR FILING DATE: 2003-08-26
;; PRIOR APPLICATION NUMBER: DE 102 39 073.8
;; PRIOR FILING DATE: 2002-08-26
;; NUMBER OF SEQ ID NOS: 69
;; SOFTWARE: PatentIn version 3.3
;; SEQ ID NO 61
;; LENGTH: 5860
;; TYPE: DNA
;; ORGANISM: Artificial sequence
;; FEATURE:
;; OTHER INFORMATION: Description of the artificial sequence: plasmid
US-10-525-674-61

Query Match 22.0%; Score 1136.4; DB 8; Length 5860;
Best Local Similarity 99.9%; Pred. No. 1.3e-259;
Matches 1137; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3405 GGGCTGCTAAAGGAAAGCGGAAACACCTAGAAAGCCAGTCCGAGAAACGGTGTGACCCCG 3464
DB 1652 GGGCTGCTAAAGGAAAGCGGAAACACCTAGAAAGCCAGTCCGAGAAACGGTGTGACCCCG 1711

QY 3465 GATGAATGTACGCTACTGGGCTATCTGGAACAGGAAACGCAAGCGCAAGAAAGCA 3524
DB 1712 GATGAATGTACGCTACTGGGCTATCTGGAACAGGAAACGCAAGCGCAAGAAAGCA 1771

QY 3525 GGTAGCTTCAGTGGGCTTACATGCGGATAGTACAGTGGCGGTTTTATGACAGCAAG 3584
DB 1772 GGTAGCTTCAGTGGGCTTACATGCGGATAGTACAGTGGCGGTTTTATGACAGCAAG 1831

QY 3585 CGAACCGGAATGCCAGCTGGGCGCCCTCTGTAAGTTGGGAAGCCCTGCAAGTAAA 3644

| | | | |
|----|------|---|------|
| Db | 2372 | CGACCAACGAGGGAACATCGATCGAGGAGACAGTACTCGANTGBAGCCGGTCTTGT | 2431 |
| Qy | 4185 | CGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCGCCGAGCCGAACCTGTTGCCAG | 4244 |
| Db | 2432 | CGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCGCCGAGCCGAACCTGTTGCCAG | 2491 |
| Qy | 4245 | GCTCAAGGCGCGCATGCCGACGGCGAGGATCTCGTGTGACCCATGGCGATGCCCTGCTT | 4304 |
| Db | 2492 | GCTCAAGGCGCGCATGCCGACGGCGAGGATCTCGTGTGACCCATGGCGATGCCCTGCTT | 2551 |
| Qy | 4305 | GCCGAATATCATGGTGGAAATGGCGCTTTTCTGATTCATCGACTGTGGCCGCTGG | 4364 |
| Db | 2552 | GCCGAATATCATGGTGGAAATGGCGCTTTTCTGATTCATCGACTGTGGCCGCTGG | 2611 |
| Qy | 4365 | TGTGGGACCCCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGG | 4424 |
| Db | 2612 | TGTGGGACCCCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGG | 2671 |
| Qy | 4425 | CGGCGAATGGGCTGACCGCTTCTCGTGTGCTTTACGGTATCGCCGCTCCGATTCGACGG | 4484 |
| Db | 2672 | CGGCGAATGGGCTGACCGCTTCTCGTGTGCTTTACGGTATCGCCGCTCCGATTCGACGG | 2731 |
| Qy | 4485 | CATCGCTTCTATCGCTTCTTTGACGAGTTCTTCTGAGCGGACTCTGGGGTTGATA | 4542 |
| Db | 2732 | CATCGCTTCTATCGCTTCTTTGACGAGTTCTTCTGAGCGGACTCTGGGGTTGATA | 2789 |

Search completed: February 27, 2006, 14:15:00
Job time : 1264 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 27, 2006, 08:34:41 ; Search time 830 Seconds
(without alignments)
11087.273 Million cell updates/sec

Title: US-10-618-570-1
Perfect score: 5177
Sequence: 1 ttgtaagaccaccaccgta.....agcgggggttttctatttgg 5177

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues
Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/1_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/H_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq.*
7: /cgn2_6/ptodata/1/ina/PP_COMB.seq.*
8: /cgn2_6/ptodata/1/ina/RE_COMB.seq.*
9: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|-------------------|
| 1 | 2078.8 | 40.2 | 7699 | 3 | US-09-645-004-1 |
| 2 | 1661.6 | 32.1 | 8316 | 2 | US-07-753-520B-4 |
| 3 | 1629.4 | 31.5 | 5594 | 3 | US-09-380-190A-29 |
| 4 | 1588.6 | 30.7 | 5130 | 3 | US-09-897-511A-9 |
| 5 | 1486 | 28.7 | 6365 | 2 | US-08-352-990-1 |
| 6 | 1486 | 28.7 | 6620 | 3 | US-07-786-531B-3 |
| 7 | 1486 | 28.7 | 9115 | 2 | US-07-753-520B-3 |
| 8 | 1455.2 | 28.1 | 5865 | 3 | US-08-654-737B-3 |
| 9 | 1448.4 | 28.0 | 8591 | 2 | US-08-462-859A-6 |
| 10 | 1448.4 | 28.0 | 8591 | 2 | US-08-462-859A-8 |
| 11 | 1448.4 | 28.0 | 8591 | 2 | US-08-123-659A-6 |
| 12 | 1448.4 | 28.0 | 8591 | 2 | US-08-123-659A-8 |
| 13 | 1448.4 | 28.0 | 8591 | 2 | US-08-464-247A-6 |
| 14 | 1448.4 | 28.0 | 8591 | 2 | US-08-464-247A-8 |
| 15 | 1448.4 | 28.0 | 8591 | 2 | US-08-464-248A-6 |
| 16 | 1448.4 | 28.0 | 8591 | 2 | US-08-464-248A-8 |
| 17 | 1434 | 27.7 | 6700 | 3 | US-09-759-152A-1 |
| 18 | 1434 | 27.7 | 8518 | 3 | US-09-654-449-3 |
| 19 | 1434 | 27.7 | 8518 | 3 | US-09-759-152A-4 |
| 20 | 1434 | 27.7 | 8518 | 3 | US-09-654-499-3 |
| 21 | 1429.2 | 27.6 | 6700 | 3 | US-09-654-449-1 |
| 22 | 1429.2 | 27.6 | 6700 | 3 | US-09-654-499-1 |
| 23 | 1425 | 27.5 | 5689 | 3 | US-09-508-516-1 |
| 24 | 1171.8 | 22.6 | 3176 | 9 | 5212080-1 |

| | | | | | | |
|------|--------|------|------|---|--------------------|-------------------|
| C 25 | 1170.4 | 22.6 | 3974 | 3 | US-09-026-343-33 | Sequence 33, Appl |
| C 26 | 1170.4 | 22.6 | 3974 | 3 | US-09-042-105-16 | Sequence 16, Appl |
| C 27 | 1170.4 | 22.6 | 3974 | 3 | US-09-044-856A-7 | Sequence 7, Appl |
| C 28 | 1170.4 | 22.6 | 3974 | 3 | US-09-023-082A-147 | Sequence 147, App |
| C 29 | 1170.4 | 22.6 | 3974 | 3 | US-09-044-855A-7 | Sequence 7, Appl |
| C 30 | 1170.4 | 22.6 | 3974 | 3 | US-09-078-670-4 | Sequence 4, Appl |
| C 31 | 1170.4 | 22.6 | 3974 | 3 | US-09-026-408-14 | Sequence 14, Appl |
| C 32 | 1170.4 | 22.6 | 3974 | 3 | US-09-362-871-33 | Sequence 33, Appl |
| C 33 | 1170.4 | 22.6 | 3974 | 3 | US-09-627-154-4 | Sequence 4, Appl |
| C 34 | 1170.4 | 22.6 | 3974 | 3 | US-09-027-287-50 | Sequence 50, Appl |
| C 35 | 1170.4 | 22.6 | 3974 | 3 | US-09-437-602-4 | Sequence 4, Appl |
| C 36 | 1170.4 | 22.6 | 3974 | 3 | US-09-252-656B-50 | Sequence 50, Appl |
| C 37 | 1170.4 | 22.6 | 3974 | 3 | US-09-248-998-147 | Sequence 147, App |
| C 38 | 1170.4 | 22.6 | 3974 | 3 | US-09-105-470B-25 | Sequence 25, Appl |
| C 39 | 1170.4 | 22.6 | 3974 | 3 | US-09-523-323-50 | Sequence 50, Appl |
| C 40 | 1170.4 | 22.6 | 3974 | 3 | US-09-902-684-14 | Sequence 14, Appl |
| C 41 | 1170.4 | 22.6 | 3974 | 3 | US-09-610-651-147 | Sequence 147, App |
| C 42 | 1170.4 | 22.6 | 3974 | 3 | US-10-084-488-16 | Sequence 16, Appl |
| C 43 | 1170.4 | 22.6 | 3974 | 3 | US-10-141-965-7 | Sequence 7, Appl |
| C 44 | 1170.4 | 22.6 | 3974 | 3 | US-10-628-395-14 | Sequence 14, Appl |
| C 45 | 1170.4 | 22.6 | 3974 | 3 | US-09-345-373-147 | Sequence 147, App |

ALIGNMENTS

RESULT 1
US-09-645-004-1
; Sequence 1, Application US/09645004
; Patent No. 6608238
; GENERAL INFORMATION:
; APPLICANT: Gavora, Jan S.
; APPLICANT: Falconer, Marcia M.
; APPLICANT: Nguyen, Thuy H.
; APPLICANT: Benkel, Bernhard F.
; TITLE OF INVENTION: Trans-Somatics With Gene Transfer Into Mammary Epithelial Cells
; FILE REFERENCE: GALA-06402
; CURRENT APPLICATION NUMBER: US/09/645.004
; CURRENT FILING DATE: 2000-08-24
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Plasmid pLAIN
US-09-645-004-1

| | | | | |
|-----------------------|--------------|--|------------------|---------------------|
| Query Match | 40.2% | Score 2078.8; | DB 3; | Length 7699; |
| Best Local Similarity | 64.3% | Pred. No. 0; | | |
| Matches 3339; | Conservative | 0; | Mismatches 1814; | Indels 38; Gaps 14; |
| Qy | 1 | TTTGAAGAGCCCCACCCGTAGGTGGCAAGCTTAAGTAAACGCCACTTTGCAAGGCAT | 60 | |
| Db | 175 | TTTGAAGAGCCCCACCCGTAGGTGGCAAGCTTAAGTAAACGCCACTTTGCAAGGCAT | 234 | |
| Qy | 61 | GGAAATATACATACCTGAGATAGAAAAGTTTCAAGTCAAGTCAAGCAAGAAACAGC | 120 | |
| Db | 235 | GGAAATATACATACCTGAGATAGAAAAGTTTCAAGTCAAGTCAAGCAAGAAACAGC | 294 | |
| Qy | 121 | TGAATACCAACAGGATATCTGTGTAAGGGTTCCTGCCCGGCTCAGGGCCCAAGAAC | 180 | |
| Db | 295 | TGAATACCAACAGGATATCTGTGTAAGGGTTCCTGCCCGGCTCAGGGCCCAAGAAC | 354 | |
| Qy | 181 | GATGAGACAGCTGAGTGAATCGGCCCAACAGGATATCTGTGTAAGGAGTTCTTCTCCCGG | 240 | |
| Db | 355 | GATGAGACAGCTGAGTGAATCGGCCCAACAGGATATCTGTGTAAGGAGTTCTTCTCCCGG | 414 | |
| Qy | 241 | CTCGGGCCCAAGAACAGATGTTCCCGGATCGGTCAGGCCTCAGCAGTTTCTAGTGA | 300 | |
| Db | 415 | CTCGGGCCCAAGAACAGATGTTCCCGGATCGGTCAGGCCTCAGCAGTTTCTAGTGA | 474 | |

QY 301 TCATCAGATGTTTCCAGGGTCCCAAGGACCTGAAATGACCCCTGTACCTTATTGAAAC 360
DB |||||
QY 475 TCATCAGATGTTTCCAGGGTCCCAAGGACCTGAAATGACCCCTGTACCTTATTGAAAC 534
DB |||||
QY 361 TAAACCAATCAGTTCTGCTGTTCTGTTGCGGCGCTTCCGCTTCCGAGCTCAATAAA 420
DB |||||
QY 535 TAAACCAATCAGTTCTGCTGTTCTGTTGCGGCGCTTCCGCTTCCGAGCTCAATAAA 594
DB |||||
QY 421 AGAGCCCAACCCCTCACTCGGGCGCGGACGTTCCGATAGACGTCGTCGCCGGGTAC 480
DB |||||
QY 595 AGAGCCCAACCCCTCACTCGGGCGCGGACGTTCCGATAGACGTCGTCGCCGGGTAC 654
DB |||||
QY 481 CCGTATTCCTCAATAAAGCCTCTTGCTGTTGTCATCCGAATCGTGTCTCGCTGTTCCCTTG 540
DB |||||
QY 655 CCGTATTCCTCAATAAAGCCTCTTGCTGTTGTCATCCGAATCGTGTCTCGCTGTTCCCTTG 714
DB |||||
QY 541 GGAGGGTCTCTCTCAGTGATTTGACTACCCAGACGGGGCTTTTCATTTGGGGCTCGT 600
DB |||||
QY 715 GGAGGGTCTCTCTCAGTGATTTGACTACCCAGACGGGGCTTTTCATTTGGGGCTCGT 774
DB |||||
QY 601 CCGGGATTGGAGACCCCTGCCAGGGACCAACCGACCCACCGGAGGTAAAGCTGGCC 660
DB |||||
QY 775 CCGGGATTGGAGACCCCTGCCAGGGACCAACCGACCCACCGGAGGTAAAGCTGGCC 834
DB |||||
QY 661 AGCAACTTATCTGTCTCTCGGATTGTCTAGTGTCTATGTTGATGTTATGCGCTCGC 720
DB |||||
QY 835 AGCAACTTATCTGTCTCTCGGATTGTCTAGTGTCTATGTTGATGTTATGCGCTCGC 894
DB |||||
QY 721 TCTGTACTAGTTAGTAACTAGCTCTGTATCTGCGGACCCGCTGGTGGAACTGACGAGTT 780
DB |||||
QY 895 TCTGTACTAGTTAGTAACTAGCTCTGTATCTGCGGACCCGCTGGTGGAACTGACGAGTT 954
DB |||||
QY 781 CTGACACCCCGCCCAACCCCTGGGAGAGTCCGAGGACCTTGGGGCCGCTTTTGTGG 840
DB |||||
QY 955 CTGAAACCCCGCCCAACCCCTGGGAGAGTCCGAGGACCTTGGGGCCGCTTTTGTGG 1014
DB |||||
QY 841 CCGGACCTGAGGAAGGAGTGCATGTGGAATCCGACCCCGTCAGGATATGTGTTCTGGT 900
DB |||||
QY 1015 CCGGACCTGAGGAAGGAGTGCATGTGGAATCCGACCCCGTCAGGATATGTGTTCTGGT 1074
DB |||||
QY 901 AGGAGACGAGAACCTTAAACAGTTCCCGCTCCGCTCGAATTTTGTCTTTCGGTTGGAA 960
DB |||||
QY 1075 AGGAGACGAGAACCTTAAACAGTTCCCGCTCCGCTCGAATTTTGTCTTTCGGTTGGAA 1134
DB |||||
QY 961 CCGAGCCCGCGCTCTGTCTGCTCGACGACGTTGGGCTGCAGTCTGATAGGAA 1020
DB |||||
QY 1135 CCGAAGCCCGCGCTCTGTCTGCTCGACGCTGCAGCATCGTTCTGTGTTGTCTGTCT 1194
DB |||||
QY 1021 TCAATTCGGCACGAGTAAATCGGTGCTGCGCTCTTTAGGACATATGAAGTATGGCACAGT 1080
DB |||||
QY 1195 GACTGTGTTTCTGTATTTGTTCTGAAATTTAGGGCCAGACTGTTACCACTCCCTTAAGTTT 1254
DB |||||
QY 1081 GGGATGATTTTCTGATCAGCAAGAGGACACTGACAGCTGTACAGAGTCTGTGAAGTTTCG 1140
DB |||||
QY 1255 GACCTTAGTCTACTGGAAGAGTGTGAGCGGATCGCTCACACCACTCGGTAGATGTCAA 1314
DB |||||
QY 1141 ATGCTCGCTCAGTGACAGCTTTGCTCTCTCCCATCTTAAATGGCCCACTCTTCAAG 1200
DB |||||
QY 1315 GAAAGAGCTTTGGGTTACTCTTCTGCTCTGCAAGATGGCCACCTTTAAGCTCGATGGCC 1374
DB |||||
QY 1201 AGAGGATGAAGTCTTATAAACTGACATGACCTCTTATCTCATGTGTTTGTAGTTTC 1260
DB |||||
QY 1375 GCGAGACGGCACCTT-TAAACGAGACCTCATCACCAGGTTAAGNATCAAGTCTT--TTT 1431
DB |||||
QY 1261 TCGTCCCATCATTTGGCATAGTGGCAGCTCAGCTCTCTGAAATGGGAAACGAAAGATTGCA 1320
DB |||||
QY 1432 ACCTGGCCCGCATGAGACCCAGACCAAGTCCCTTACATCGTACCTGGGAAGCCTTGGC 1491
DB |||||
QY 1321 CCGTGTGCTCAGTTAATTCAGATATATCTCCAGTCCGAGGCAAGGAAATGCGAGTG 1380
DB |||||
QY 1492 TTTTGACCCCTCTCCCTGGGTCAAGCCCT-----TTGTACACCTTAAGCCTCCGCTCT 1546
DB |||||
QY 1381 AAGATGAATGAGATTTTCGAGAGCTGTGATGGAAAGCAGCATGAGCAATGGAAGACGAA 1440

DB |||||
QY 1547 CTTCTCCATCGGCCCGCTCTCTCCCTTGAACTCTCTCGTTCCAGCCCGCTCGATCC 1606
DB |||||
QY 1441 TCCAGTATCTTTTTCAGATAATGAAGCCAAATCTCTCTAGATGCTTAAGAAATTTCCAAATTTCA 1500
DB |||||
QY 1607 TCCCTTTATCAGGCCCTCACTCTCTCTAGGCCCGGAAATTCGTTAACTCGACATGAA 1666
DB |||||
QY 1501 GCATAACAACCTGATCAAGAGATTTAAATGATGTTCTTTTCCAGCTAAATTCCTTACTTCTCT 1560
DB |||||
QY 1667 GTCCTTCTCTCTCTCGCAGCTGTGCGGCTTTGCTGGGCACAGTACAATCCCAACACTCAG 1726
DB |||||
QY 1561 CCATCCAGG--AACATGAGAATATCATAGGGGATATCTCTCAAGTCAATTTAGTAGSTCTGAA 1618
DB |||||
QY 1727 GCTGGGAGGACATCTATCGTGCACTCTCTTTGAATGGCGCTGGGCCGACATTCGCACTGAG 1786
DB |||||
QY 1619 CACCAAGTACTTGAATTTGACAGTTTCACTAATGATCTGAGGCTGAAGGATTTGGGAACTTCTCAGAC 1678
DB |||||
QY 1787 TGCGAACACTATTATTAGCTCTTAATGGGTTTGGAGGAGTTTCAAGGTTTCTCTCTCCAAATGAA 1846
DB |||||
QY 1679 TGCATTTTAAACAACAAGAGGAGATGCGTAAATTTAGAGGAGGTATATACAATGCATCAGC 1738
DB |||||
QY 1847 AACATGTCATTACTAATCCGAACAGGCCCTGCTGGTGGAAAGATACCAAGCCCATCAGCTAC 1906
DB |||||
QY 1739 AGAAATTAAGTCTCTAGATGAAAAAACAAGTATATTTTGGAAACAGGAAATTAAGAAGGGGAAT 1798
DB |||||
QY 1907 AAGATCTCGAGTCGATCGGGCAATGAATAATGAATTCAGAGACATGGTGACCAGATGCAAC 1966
DB |||||
QY 1799 GAAACTGTTGAAATTAATTAATGATCTGAGGCTGAAGGATTTGGGAACTTCTCTCAGAC 1858
DB |||||
QY 1967 AATGTTGGAG--TTGCTATTTATGTTGATGCTGTTGTCAATCACATGTGTGATCTATGGG 2025
DB |||||
QY 1859 ATTGAAAATATCACTTTTACTCCTCAAGTGGCCAGAAAGTGTCTCGCTGACTGGGAAATGGAC 1918
DB |||||
QY 2026 TGGCACGGGACCCCACTCAACATGTGGGAGCTATTTCAACACCCGGGACTAGAGATTTTCC 2085
DB |||||
QY 1919 CAACGATCTGGGCTCCAAACATGACCATCGGGGCTGTGAAACAGAGAGAGTGAAATTCACAGG 1978
DB |||||
QY 2086 CGCTGTCCGTACTCTGCTGGGATTTCAATGACGGCAATGTCACTCACTGCAAGTGGAGA 2145
DB |||||
QY 1979 CACTTACATCAGAGCCGTAAACAGCCACATCAATATGAGATCAAGAGTCAACACTGCTATGG 2038
DB |||||
QY 2146 CATCGAAAATTTATGGGACATGTATCAGGTCCGGGATTGCAAGTTGTCCAGCCTTCTTGA 2205
DB |||||
QY 2039 GACA-CAAAACACCATCAACAGAGGACCCAGCCCACTTTGGCTTCACTCGTCAATTTGGA 2097
DB |||||
QY 2206 TCTGGCTCTGGAGAGGACTATGTACGCTCAACAAATTTGACGCTACATGAATCACCTCAT 2265
DB |||||
QY 2098 AGTTTTCAGAGTCCACCACTGTCTTTCAGGGGCCAGTGTCTTATAGACAGGAATGGGAAGG 2157
DB |||||
QY 2266 TGATATGGGTGAGCAGGGTTCCGGATCGATGCTGCCAAGCATATGTGGCCAGGGACAT 2325
DB |||||
QY 2158 AGGTCTTGAAGACATGTGGCTGTGCGGTCAAGTGTTAATGACATTTGGTGTATGACTGGA 2217
DB |||||
QY 2326 AAGAGCGTTTCTGAGCAAACTGCACGATCTAAATPACTCAGTGGTGTTCAGCAGGAACGAA 2385
DB |||||
QY 2218 AAGCTACAGGGTC-GGCATCAACATCTTCACTCGCTCGCACACAGAGAGTGTAGTG 2276
DB |||||
QY 2386 ACCCTTTATTACCAAGAGTAAATTGACTTTGGGAGGAGAGCCAAATCAGGCACTCAGTA 2445
DB |||||
QY 2277 AGTGACCAAGGCTCCTCTG--GACTCCAGGTGAAAAGGAGATAGAGGCCCTCTCTGGAC 2333
DB |||||
QY 2446 CTTTGGGAAATGGCCGCTGACAGAAATTCAGATGTTGGTGCCTCAACTGGGACGGTGATCCG 2505
DB |||||
QY 2334 AAAATGTTATACAGGCTTTCCAGGTCTAATAGGTACTCAGAGTCTTAAAGGTGTATCGGG 2393
DB |||||
QY 2506 GAAGTGAATGGAGAGAGATGGCTTACTTAAAGAACTGGGGAGAGGCTGGGGCTTGT 2565
DB |||||
QY 2394 GGGATCTCTGTTTACCTGGGTTCCAGGATTTCCAGGAC--CAATGGGGAAGACCCGGA 2451
DB |||||
QY 2566 GCCTTCTGACAGAGCCCTGTGTGTTTGTGGATTAACACGACAAACCAAGCGGGGGACGGGGC 2625
DB |||||
QY 2452 AGCCAGGACTTTAATGGAACAAAAGGCCAGAGGGAGAAAAAGGAGGTGGAAAGCATGCAAA 2511
DB |||||

Db 2626 AGCGGAGCTCCATTCTTCTGCGATGCCAGGCTTTTATAAATGGCGGTGGTTT 2685
Qy 2512 GACAATCTAATACAGTCCGATCTGTGGTGGCAGCGCCCTCAAGAGGACAGATGGAGA 2571
Db 2686 CATGCTCGCTCATCCGTACGGGTTTCAACGGCTGATGTCAAGTTATCTGTGGCCAAGATA 2745
Qy 2572 TTTTTCAGAGGCGCAGTGGGTACGGTGTGTGACGCGCTGGGAACTGGTGGAGAC 2631
Db 2746 TTTTTCAGAGGCGCAGTGGGTACGGTGTGTGACGCGCTGGGAACTGGTGGAGAC 2805
Qy 2632 T----GGTGTGTGCGAGGCTTGGGATACAAAGGTGTCAAAGTGTGCATAAGCGAGCT 2687
Db 2806 GACGAAGTCCGTACCAATCAACGCAGACACTACCTGTGGCAATGACTGGGTCTGGGAACA 2865
Qy 2688 TATTTTGAAGAGTACGGGTCCAATATGGCTGAATGAATTTTGTTCGGGAAGAG 2747
Db 2866 TCGCTGGCGACAAATAAGG----AACATGGTTATCTTCCGTTAATCTCCGTAAATGTGTAGACGTCAGC 2921
Qy 2748 TCATCCATTGAAGAGTGCAGATTTAGACAGTGGGGTGTGAGAGCTGTTTCCACAGCGAA 2807
Db 2922 CTTTCTCAAACTGGTGGACAAACGGGAGCAATCAAGTAGCTTTTCGGTCCGCGGCGACAGAG 2981
Qy 2808 GATGCTGGGGTCACTTTGCACTACATAATATGTCATATTTTCAATTCACATTTTTTAA 2867
Db 2982 GCTTCATGTCTTAAATGATGACTGGTATATGAATGTGATTTGCAACTGGTCTGC 3041
Qy 2868 CTGTATAAAGTGATTTTTTCTTGTCTTCACTAAATACAGTTAAATTAATTTAAGA 2927
Db 3042 CTGCTGGAACCTACTGCGATGTATTTCTGACAAAGAGGAGCGAGTGTACTGAA 3101
Qy 2928 AACTAAGATTTTATCCAGAAAAGGATATTTAAATCACTGGATAACATATAAA 2987
Db 3102 AGCAGGTGTACGTTCTCGATGGAAAGGCCAATTTCCAGATTAGTAACAGCGATGAAG 3161
Qy 2988 TAGCTTCATATTTGCTTCAATATACAGAAACCAATTTCAATCTCTAGTGTTTTAAAGTGGC 3047
Db 3162 ATCAATTTGTGCAATTCAGTTGATGCAAGTATTAAGCTTCGAGGATCCACTAGTAAC 3221
Qy 3048 TCGTGCCGAATTTGATCCCTCAGGATATAGTATTTTCGCTTTGCATAGGAGGGGAAA 3107
Db 3222 GGCGCCAGTGTCTGGAATTCG-----CTTGTGCAATCTAGGGCGGCAATTCGCC 3276
Qy 3108 TGTAGTCTTATGCAATCTCTTGTAGTCTTGCAACATGTAACGATGATGATGCAACATG 3167
Db 3277 CCTCTCCCGCCCGCCCTTAAAGTTACTTGGCGAAGCGCTTGGAAATAAGGCGCGTGTGTG 3336
Qy 3168 CTTTACAGGAGAGAAAAGCACCGTGCATGCCGATTTGGTGAAGTAAGTGTGACGATC 3227
Db 3337 TTTGTCTATATGTGATTTTCCACCATATGTCCGCTTTTGGCAA-----TGTGAGGGCC 3390
Qy 3228 GTGCTTTATTAGGAAGGCAACAGACGGGTCTGACATGATTTGGAGCAACCACTGAATTC 3287
Db 3391 CGAAACCTGGCCCTGTCTTCTTACGAGCATTTCTAGGGGTCTTTCCCTCTCGCCAA 3450
Qy 3288 GCATTCAGAGATATGTAATTAAGTCCCTAGCTCGATPACAGCAACGCCATTTGACCAT 3347
Db 3451 GGAATGCAAGTCTGTGTAATGTCTGAAAGAACAGTCTCTCTGGAAGCTTCTTGAAGA 3510
Qy 3348 TCACCATTTGGTGTGACCTTCAAGCTTACGCTGCCGAGCACTCAGGGCGCAAGGG 3407
Db 3511 CAACAACAGTCTGTAGCGACCTTTTTCAGGAGCGGAAACCCCGCCAGGCGACAGGTGC 3570
Qy 3408 CTGCTAAAGGAAGGGGAACAGTGAAGAGCCAGTCCGAGAAAGCGGTGCTGACCCCGAT 3467
Db 3571 CTCTGGGCCAAAGCCAGTGTATAGATACACTGCAAGGGCGGCAACCCAGTGC 3630
Qy 3468 GAATGTACGATCTGGGTATCTGACAC-AGGGAAGCGAAGCGCAAGAGAACAGG 3526
Db 3631 CACGTTGTGAGTGTGATAGTTGTGGAAGAGTCAATGCGCTCTCTCAAGCGTGTCAAC 3690
Qy 3527 TAGCTTGCAGTGGGCTTACATGGCGATAGTAGTGGCGGCTTTTATGGACAGCAAGCG 3586
Db 3691 AAGGGGCTGAAGGATGCGCCAGAGGTACCCCAATTTGTATGGGAATCTGATCTGGGGCGCTCG 3750

Qy 3587 AACCGGAATTCAGCTGGGGCGCCCTCTCTGTATAGGTTGGGAAGCCCTGCAAAAGTAAACT 3646
Db 3751 GTGCACATGCTTTACATGTGTGTTAGTCAGGTTAAAAAGCTCTAGGCCCCCGCAACCAC 3810
Qy 3647 GGATGGCTTTCTTGGCCCGCAGGATCTGATGCGCGCAGGGATCAAGATCTGATCAAGAGA 3706
Db 3811 GGGGACGTGGTTTCTCTTTGAAAAACACAGATATAGCTTGGCCACAACCCGGGATAATTC 3870
Qy 3707 CAGGATGAGGATCGTTTTCGATGATTAACAAGATGATGACGCGAGGTTCTCCGGCGC 3766
Db 3871 CTGCAGCAATATGGGATCGCCCATTCGAACAGATGATGCAAGCGAGGTTCTCCGGCGC 3930
Qy 3767 CTTGGGTGGAGAGGCTATTTCCGCTATGACTGGGCAACAAGCAATCGGTGCTCTGATG 3826
Db 3931 CTTGGGTGGAGAGGCTATTTCCGCTATGACTGGGCAACAAGCAATCGGTGCTCTGATG 3990
Qy 3827 CCGCGGTGTTCCGCTCTCAGCGAGGGCGCGGCTTCTTTTGTCAAGACCGACCTGT 3886
Db 3991 CCGCGGTGTTCCGCTCTCAGCGAGGGCGCGGCTTCTTTTGTCAAGACCGACCTGT 4050
Qy 3887 CCGGTGCTCCGATGAATGAATCTGAGGAGGCGCGGCTATCTGCTGGCTGGCCACAGCG 3946
Db 4051 CCGGTGCTCCGATGAATGAATCTGAGGAGGCGCGGCTATCTGCTGGCTGGCCACAGCG 4110
Qy 3947 GCCTTCTTGGCAGCTGTGCTGACGCTTGTCTGAAAGCGGGAAGGAGCTTGGCTGTAT 4006
Db 4111 GCGTTCCTTGGCAGCTGTGCTGACGCTTGTCTGAAAGCGGGAAGGAGCTTGGCTGTAT 4170
Qy 4007 TGGCGAAGTGGCGGGCAGGATCTCTGTCTCATCTCACTTGTCTCTCCGCGAAGAGTAT 4066
Db 4171 TGGCGAAGTGGCGGGCAGGATCTCTGTCTCATCTCACTTGTCTCTCCGCGAAGAGTAT 4230
Qy 4067 CCATCATGGCTGATGCAATCGCGGCTGCAATCGCTTGTGATCGGCTACCTGCCCATTCG 4126
Db 4231 CCATCATGGCTGATGCAATCGCGGCTGCAATCGCTTGTGATCGGCTACCTGCCCATTCG 4290
Qy 4127 ACCAACAGGCAAACTCGCATCGAGGAGCAGCTACTCGGATGGAAGCGGCTTGTGTCG 4186
Db 4291 ACCAACAGGCAAACTCGCATCGAGGAGCAGCTACTCGGATGGAAGCGGCTTGTGTCG 4350
Qy 4187 ATCAGGATGATCTGGAGCAAGAGCATCAGGGGCTCGCGCGCAGCGCAACTGTTGCCAGGC 4246
Db 4351 ATCAGGATGATCTGGAGCAAGAGCATCAGGGGCTCGCGCGCAGCGCAACTGTTGCCAGGC 4410
Qy 4247 TCAAGGCGCATGCCCGAGCGGAGGATCTCTGCTGTGACCCATGGCGATGCCCTGCTTCG 4306
Db 4411 TCAAGGCGCATGCCCGAGCGGAGTCTCTGCTGTGACCCATGGCGATGCCCTGCTTCG 4470
Qy 4307 CGAATATCATGTTGGAAAAATGGCGCTTTTCTGATTCATCGACTGTGGCGCGCTGGGTG 4366
Db 4471 CGAATATCATGTTGGAAAAATGGCGCTTTTCTGATTCATCGACTGTGGCGCGCTGGGTG 4530
Qy 4367 TGGCGGACCGCTATCAGGACATAGCGTTGGCTACCGCTGATATTTGCTGAAGAGCTTGGCG 4426
Db 4531 TGGCGGACCGCTATCAGGACATAGCGTTGGCTACCGCTGATATTTGCTGAAGAGCTTGGCG 4590
Qy 4427 GCGAATGGGCTGACCGCTCTCTGCTGCTTTACGCTATCGCGCTCCGATTTGCGAGCGCA 4486
Db 4591 GCGAATGGGCTGACCGCTCTCTGCTGCTTTACGCTATCGCGCTCCGATTTGCGAGCGCA 4650
Qy 4487 TCGCTTCTATCGCTTCTTGAAGAGTCTTTCTGAGCGGAGCTCTGGGGTTCGATAAAAT 4546
Db 4651 TCGCTTCTATCGCTTCTTGAAGAGTCTTTCTGAGCGGAGCTCTGGGGTTCGATAAAAT 4710
Qy 4547 AAAAGATTTTATTTAGTCTCCAGAAAAAGGGGGAATGAAGACCCCACTGTAGGTTTG 4606
Db 4711 AAAAGATTTTATTTAGTCTCCAGAAAAAGGGGGAATGAAGACCCCACTGTAGGTTTG 4770
Qy 4607 GCAAGCTAGCTTAAGTAACCCCAATTTTGAAGCATCGGAATATACATACTGAGATAG 4666
Db 4771 GCAAGCTAGCTTAAGTAACCCCAATTTTGAAGCATCGGAATATACATACTGAGATAG 4830

QY 4667 AGAAGTTTCAGATCAAGGTGAGGAGCAGATGGAACAGCTGAATATATGGCCAAACAGGATAT 4726
DB 4831 AGAAGTTTCAGATCAAGGTGAGGAGCAGATGGAACAGCTGAATATATGGCCAAACAGGATAT 4890
QY 4727 CTGTGGTAAAGCAGTTCTGCGCCCGGCTCAGGGCCAGAACAGATGGAACAGCTGAATATG 4786
DB 4891 CTGTGGTAAAGCAGTTCTGCGCCCGGCTCAGGGCCAGAACAGATGGAACAGCTGAATATG 4950
QY 4787 GGCACAAACAGGATATCTGTGTGTAAGCAGTTCTGCGCCCGGCTCAGGGCCAGAACAGATG 4846
DB 4951 GGCACAAACAGGATATCTGTGTGTAAGCAGTTCTGCGCCCGGCTCAGGGCCAGAACAGATG 5010
QY 4847 GTCCCCAGATCGGTTCAGACCCCTCAGCAGTTTCTAGAGAACCATCAGATGTTTCCAGGGT 4906
DB 5011 GTCCCCAGATCGGTTCAGACCCCTCAGCAGTTTCTAGAGAACCATCAGATGTTTCCAGGGT 5070
QY 4907 GCCCAAGGACCTGAAATGACCCCTGTCCTTATTTGAACCTAACCAATCAGTTCGCTTCTC 4966
DB 5071 GCCCAAGGACCTGAAATGACCCCTGTCCTTATTTGAACCTAACCAATCAGTTCGCTTCTC 5130
QY 4967 GCTTCTGTTCGCGGCTTCTGCTCCCGAGCTCAATTAAGAGCCCAACCCCTCACTC 5026
DB 5131 GCTTCTGTTCGCGGCTTCTGCTCCCGAGCTCAATTAAGAGCCCAACCCCTCACTC 5190
QY 5027 GGGCGCCAGTCTCCGATTGACTGAGTCGCGCGGTACCCGTGTATCCAATAAACCCCTC 5086
DB 5191 GGGCGCCAGTCTCCGATTGACTGAGTCGCGCGGTACCCGTGTATCCAATAAACCCCTC 5250
QY 5087 TTGCAGTTGCAATCCGACTGTGTGCTCTGCTGTTCTTGGAGGGTCTCCTCTGAGTGATT 5146
DB 5251 TTGCAGTTGCAATCCGACTGTGTGCTCTGCTGTTCTTGGAGGGTCTCCTCTGAGTGATT 5310
QY 5147 GACTACCCGTGAGCGGGGCTCTTTCATTGG 5177
DB 5311 GACTACCCGTGAGCGGGGCTCTTTCATTGG 5341

RESULT 2

US-07-753-520B-4
; Sequence 4, Application US/07753520B
; Patent No. 5352595
; GENERAL INFORMATION:
; APPLICANT: Tapscott, J.; Weintraub, H.M.; Palmer, T.D.
; TITLE OF INVENTION: "MYOD REGULATORY REGION"
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Christensen, O'Connor, Johnson and Kindness
; STREET: 2800 Pacific First Center, 1420 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98101-2347
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette-5.25 inch, 1.2Mb storage
; COMPUTER: IBM PC/386 Compatible
; OPERATING SYSTEM: MS-DOS 4.01
; SOFTWARE: Word for Windows-t
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07753,520B
; FILING DATE: 19910903
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: none
; FILING DATE: none
; ATTORNEY/AGENT INFORMATION:
; NAME: Sundemo, John, S.
; REGISTRATION NUMBER: 34,446
; REFERENCE/DOCKET NUMBER: FHCR-1-5789
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 1-206-682-8100; 1-206-224-0727 (direct)
; TELEFAX: 1-206-224-0779
; TELEX: 4938023
; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 8316 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Other; plasmid DNA
; DESCRIPTION: pLHDMN-NSA; 5'LTR (position 1-1159); Y+ (position 1160-1640); HisD (position 1641-2928); Myo-D NSA Apal fragment (position 2929-4389); driving neo (position 4390-5259); with 3'LTR (position 5260-5964); Figures 8A-8C.
; US-07-753-520B-4

Query Match 32.1%; Score 1661.6; DB 2; Length 8316;
Best Local Similarity 61.1%; Pred. No. 0;
Matches 3536; Conservative 0; Mismatches 1641; Indels 607; Gaps 29;
QY 1 TTTGAAAGACCCCAACCGGTAGGTGGCAAGCTTAGCTTAAGTAAACGCCACTTTTGCAGGCAT 60
DB 175 TTTGAAAGACCCCAACCGGTAGGTGGCAAGCTTAGCTTAAGTAAACGCCACTTTTGCAGGCAT 234
QY 61 GGAATAATACATAACTGAGATAGAAAGTTTCAGATCAAGGTGAGGAAACAAAGAAACAGC 120
DB 235 GGAATAATACATAACTGAGATAGAAAGTTTCAGATCAAGGTGAGGAAACAAAGAAACAGC 294
QY 121 TGAATACAAACAGGATATCTGTGTAAAGCGGTTCTGCCCCCGGCTCAGGGCCAAAGAAC 180
DB 295 TGAATACAAACAGGATATCTGTGTAAAGCGGTTCTGCCCCCGGCTCAGGGCCAAAGAAC 354
QY 181 GATGAGACAGCTGAGTGATGGGCCAAACAGGATATCTGTGTAAAGCGGTTCTGCCCCCGG 240
DB 355 GATGAGACAGCTGAGTGATGGGCCAAACAGGATATCTGTGTAAAGCGGTTCTGCCCCCGG 414
QY 241 CTGGGGCCAAAGAACAGATGTCCTCCAGATGCGGTCCAGCCCTCAGCAGTTTCTAGTGAA 300
DB 415 CTGGGGCCAAAGAACAGATGTCCTCCAGATGCGGTCCAGCCCTCAGCAGTTTCTAGTGAA 474
QY 301 TCATCAGATGTTTCCAGGGTGCCCCCAAGGACCTGAAATGACCCCTGTACCTTATTTGAAC 360
DB 475 TCATCAGATGTTTCCAGGGTGCCCCCAAGGACCTGAAATGACCCCTGTACCTTATTTGAAC 534
QY 361 TAAACCAATCAGTTGCGTTCTGCTTCTGTTCGGCGGTTCCGCTCTCCGAGCTCAATAAA 420
DB 535 TAAACCAATCAGTTGCGTTCTGCTTCTGTTCGGCGGTTCCGCTCTCCGAGCTCAATAAA 594
QY 421 AGAGCCCAACCCCTCACCTCGGGCGGCAGTCTCCGATAGACTGCGTCCCGGGGTAC 480
DB 595 AGAGCCCAACCCCTCACCTCGGGCGGCAGTCTCCGATAGACTGCGTCCCGGGGTAC 654
QY 481 CCGTATTTCCCAATAAAGCCCTTCTGCTGTTTGCATCCGAATCGTGGTCTCGCTGTTCTTG 540
DB 655 CCGTATTTCCCAATAAAGCCCTTCTGCTGTTTGCATCCGAATCGTGGTCTCGCTGTTCTTG 714
QY 541 GGAGGGTCTCCTCTGAGTGATTGACTACCAACGACGGGGGTCTTTTCATTTTGGGGGTCTGT 600
DB 715 GGAGGGTCTCCTCTGAGTGATTGACTACCAACGACGGGGGTCTTTTCATTTTGGGGGTCTGT 774
QY 601 CCGGATTTGGAGACCCCTCGCCAGGACCAACGACCCCAACGACCCGAGGAGTAAGCTCGCC 660
DB 775 CCGGATTTGGAGACCCCTCGCCAGGACCAACGACCCCAACGACCCGAGGAGTAAGCTCGCC 834
QY 661 AGCAACTTATCTGTGCTGTCCGATTGCTAGTGTCTATGTTTGTATGTCGCCCTGCG 720
DB 835 AGCAACTTATCTGTGCTGTCCGATTGCTAGTGTCTATGTTTGTATGTCGCCCTGCG 894
QY 721 TCTGTACTAGTTAGCTAACTAGCTCTGTATCTGCGGACCCGCTGGTGGAACTGACGAGTT 780
DB 895 TCTGTACTAGTTAGCTAACTAGCTCTGTATCTGCGGACCCGCTGGTGGAACTGACGAGTT 954
QY 781 CTGAACACCCCGCGCAACCCCTGGAGACCTCCAGGAGCTTTGGGGGCGGTTTTTGG 840
DB 955 CTGAACACCCCGCGCAACCCCTGGAGACCTCCAGGAGCTTTGGGGGCGGTTTTTGG 1014
QY 841 CCGGACCTGAGGAAGGGAGTCGATGTGAATCCGACCCCGTACGATATGTGTCTGTGT 900

Db 1015 |||||CCCCACCTGAGAGGGAGTCGATGTGAATCCGACCCCGTCAGGATATGGTTCTGGT 1074
Qy 901 AGGAGACGAGAACCTTAAACAGTCTCCCGCTCCGCTGAAATTTTGTCTTTCGGTTGGAA 960
Db 1075 AGGAGACGAGAACCTTAAACAGTCTCCCGCTCCGCTGAAATTTTGTCTTTCGGTTGGAA 1134
Qy 961 CGAAGCCGCGGCTCTTCTGCTGACAGCAAGCTTTGGGCTGACAGTCTCTAGAGGA 1020
Db 1135 CCGAAGCCGCGGCTCTTCTGCTGACAGCAAGCTTTGGGCTGACAGTCTCTAGAGGA 1134
Qy 1021 TCAATTCGGCAGAGTAATCGGTGCTCCGCTTTTAGGACATAGAGTATGSC----- 1075
Db 1195 TTCTGTATTCTGAAAATTAGGCGCAGACTGTACCACTCCCTTAAAGTTGACCTTAG 1254
Qy 1076 -----ACAGTGGGATGACTTCTCTGATCAGCAGAGAGACACTGACAGCTGTACAG 1125
Db 1255 GTCACCTGGAAGATGTCGAGCGGATCGCTCACACCACTCGGTAGTGTCAAGAGAGAC 1314
Qy 1126 AGTCTGTGAAGTTGAGTCTCGCTCAGTGAAGCTTTGCTTCTCCCAATCTTAAAGATG 1185
Db 1315 GTTGGGTTACCTTCTGCTCTGAGAAATGCGCAACCTTTAAAGCTCGGATGGCGGAGACG 1374
Qy 1186 GCCCAACTCTTCAAGAGAGTGAAGTCTTATATAAACTGCACTGATCACCTTTATCTCA 1245
Db 1375 GCACCTTTAAACCGAGACCTCATCCAGGTTAAAGATCAAGGTCTTTTCACTCGCCGCG 1434
Qy 1246 TTGTGTTGTTAGTTCTCGTGGCCCATTTGGCATAGTGGCAGCTCAGCTCTCTGAAATGG 1305
Db 1435 ATGNACCCAGACAGAGTCCCTCATCTGATCGTGAAGCTTGGGATGCTTTGACCCCC 1494
Qy 1306 AAACGAGAAATTGACGGTGGCTCAGTTAATGAGATATATCTCAAGTCCGGAAGGCA 1365
Db 1495 CTCCTGGGTCAAGCCCTTTGTACACCTTAAGCCTCGCCCTCTCTCTCCATCCGCGCC 1554
Qy 1366 AAGMAATGGCAGTGA-----AGATG 1386
Db 1555 CGTCTCTCCCTTTGAACCTCTCTGTCGACCCCGCTCGATCTCTCTTTTCCAGGCC 1614
Qy 1387 AAATGAGATTTGAGAGCTGTGATGGAAC-----GCATGAGCAACATGGA 1432
Db 1615 TCATCTCTTCTTAGGCGCGGAATTGATCCGGACCATGAGCTTCAATACCTGTATGA 1674
Qy 1433 AAGCAGAACTCAGTATCTTTCAGATAATGAAGCCAAATCTCTAGATGCTAAGAAATTTCCA 1492
Db 1675 CTGGAACAGCTGTAGCCCTGAACAGCAGCGTGGCTGTGACGGCTCGGCGAATTTCCGC 1734
Qy 1493 AAATTTGAGATAAC-----AATGATCAAGATTTAATGATGTTCTTTTCCAGCT 1543
Db 1735 CTCTGACAGTATTACCCGACGCGTCAGCGATATTCTGGATTAATGTAATAAACGCGGTGA 1794
Qy 1544 AAATTTCTTACTTTCTCCATCCAGGACATGAGATATCATAGG----- 1589
Db 1795 CGATGCCCTCGGTGAAATACGCGTAAATTTGATATAAACAAGATGACAGCGCTACGCGT 1854
Qy 1590 -----GATATCTCAAAGTCAATTAGTAGGCTGAAACACCACTGATTTGATTTGCA 1639
Db 1855 CACCCCTGAAGATCGCGCGCGCGCTGAGCGCAAGTAAATAAACAGCGCAT 1914
Qy 1640 GTTCAGTATTAACACTGAAATGGCAGAGTCA---AGAGATGCAATTTAAACAAAGAA 1696
Db 1915 GACCGCTCCCTCAAAAATAATTGAACCGTTCCATTCGCGCAGAGCGCTACCGCTCTGTA 1974
Qy 1697 GGAGATGCGTAAATTAGAGGAGCGTATATACATCATCAGCAGAAATTAAGTCTCTAGA 1756
Db 1975 TGTGAAACCCAGCCAGCGCGTGTGCGCAGAGGTTACGCGTCCCGCTCTGCTGTGCG 2034
Qy 1757 T---GAAAAAACAAGTATATTTGGAAACAGAAATAAAGGGAAATGAACCTGTTGAATAA 1813
Db 2035 TCTGTATATTCCGCGGCTCGGCTCTCGCTCTTCTCAACGCTGTGATGCTGGCAGGCC 2094
Qy 1814 TATCTAATGATCTGAGGCTGAAGGATTTGGGAACATTTCTCAGACATTTGAATAATATCAC 1873

Db 2095 GGCGCGCATTTGCGGGATGCCAGAAAGTGGTTCTGTGCTCGCGCGGCCCATCGCTGATGA 2154
Qy 1874 TTATCTCAAAGTGCAGAAAGTGTCTCGTACATGGGAAATGGACCAACG----- 1923
Db 2155 AATCTCTATGCGCGCAACTGTTGTGGCGTGCAGGAAATCTTTAACGTGCGCGCGCGCA 2214
Qy 1924 -----ATCTGGGCTCCAACATGACCATTCGGGCTGTGAAACAGCAGAGTGAATTCACAGG 1978
Db 2215 GCGGATTTCCGCTCTGCGCTTTCCGACAGCGAGTCCGCTACCGAAAGTGGATNAATTTTGG 2274
Qy 1979 CACCTACATCACAGCCGTAAACAGCCACATCAAAATGAGATCA--AAGAGTCAACCATGCTAT 2036
Db 2275 CCCCGGCAACGCCCTTTGTAAACGAAAGCAACAGCTCAGCTCAGCCAGCGCTCTGACGGGCG 2334
Qy 2037 GGGACACAAAACACCATC-----AAACAAGAGACCCAGCCCACTTTTGGCT 2082
Db 2335 GGCTATCGATATGCCAGCGGCGCTCTGAAGTACTGTGATCGCAGACAGCGCGCAAC 2394
Qy 2083 TCACCGTCAATTGGAAGTTTTCAGAGTCCACCATGCTTTCACGGGCC----- 2130
Db 2395 ACCGGATTTCTGCTGCTTTCGACCTGCTCTCCAGGCTGAGCAGCGCCCGGATTCACAGGT 2454
Qy 2131 -----AGTCTTTCATAGACAGGAATGSGGAAGGAGTCTCTGAAGACCAT 2173
Db 2455 GATCCTGCTGACGCTGATGCTGACATTTGCCGCAAGTGGCGGAGCGGTAGAACGTCA 2514
Qy 2174 GTGCTGCTGCGGTCAAGTGTAAATGATCTGATGATGCTGAAAGCTA--CCAGGGTCT 2231
Db 2515 ACTGGCGGAATCTGCGCGCGGACACCGCCCGCAGGCCCTGAGCGCATCTGCTGAT 2574
Qy 2232 GGCATCAACATCTTTCATCTGCTCGGCACACAGAAAGAGTGAAGTGAACAAGTCTCT 2291
Db 2575 TGTGACCAAGATTTAGCGCAGTGCCTCGCATCTCTAATCAGTATGCGCGGAACACTT 2634
Qy 2292 CCTGGACTCCAGGTGAAAAGGAGATAGAGCCCTCTTGACAA----- 2335
Db 2635 AATCATCAGACGCGCAATGCGCGGATTTTGTGGATGCGGATTCACGCGAGGCTCGGT 2694
Qy 2336 -----AATGTTATACAGGCTTTCCAGGCTTAATAGTACTTCCAGGTCTTTAAAGG 2385
Db 2695 ATTTCTCGGCGACTGTGCGCGGAATCCGCGGTGATTACGCTTCGGAACCAACCATGT 2754
Qy 2386 TGATCGGGGGATCTCTGTTTACTGAGTTCAGAGTTCAGGATTCAGGACCAATGGGGAAGA 2445
Db 2755 TTATCCGACCTATGGCTATCTGCTACTGTTTCCAGCTTGGGTTAGCGGATTTCCAGAA 2814
Qy 2446 CCGGGAAGCCAGGACTTAATGGAACAAAGG-----CCAGAAAGGA 2486
Db 2815 ACGGATGACGCTTCAGGAACCTGTGAAAGCGGGGCTTTTCCGCTCTGGCATCAACCATTTGA 2874
Qy 2487 GAAAAAGGGAGTGAAGCATGCAAAAGCAATCTAATACAGTCCGACTGTTGGTGGCAGC 2546
Db 2875 AACATTGGCGGCGAGAAACGTCTGACCGGCCATTAATAATCGCGTACCTTGGCGTAAA 2934
Qy 2547 GGCCCTCAAGAGCGCAGAGT-----GAGATTTTTCAGAAAGCCAGTGGGTTACGCTG 2600
Db 2935 CGCCCTCAAGAGCAAGCATGAGCACTGAAACACTCTCAGCGTCCGCGGATCAATTCG 2994
Qy 2601 TGTGACGACCGCTGGGAACCTGCTGGAGGACTG--GTGCTGTCAGGAGCTTGGGATACAA 2659
Db 2995 TTAACTCGAGGATCGGCCAGCCACAGATGGTATAGTATAGNATAGATTTTATTAGA 3054
Qy 2660 AGGTGTTCAAAGTGTGATAGCGAGCTTATTTTGGAAAGGTACGGGTCCAAATATGGCT 2719
Db 3055 GCATAGGAGGGAGTCAAGAGGCGAGTAGAGGCAAGAAAGAGAGAGAGGGGGAGAG 3114
Qy 2720 GAATGAAGTATTTTGTTCGGGAAAGATC--ATCCAATTGAAGAGTGCAGAAATTAGACAG 2777
Db 3115 AAGAGAGTAGAGGTAGAGGCTGGCTTGGAGCACTGAGAGAGAGCAGGGAAGGAA 3174
Qy 2778 TGGGTTGTGAGAGCTGTTTCGCAC-----GACGAAGATCTCGGGGTCACT 2823
Db 3175 TGGGAGAGAGGGNAGAAATTGGAACACAGAGAAAGCAAGAGAGATCAAGAGAGCAAGG 3234

QY 2824 TTGACCTACATAAATGCAATCATATTTTTCATTCACATTTTAACTGTTATAAAGTGATT 2883
DB AGGGGGGCAAGCAGCCCTTTTATAGTAGTACGACACACCTGGTGTTCGCGAGTAAT 3294
QY 2884 TTTTTCCTTTGCTTCACTAA--TCAGCTTAATTAATATTTAAAGAA 2929
DB GTGGGGGGGNTTAGACACAAAGCTAAACATCCACCACTAGTTTATCTGCCAGACT 3354
QY 2930 CTAAGAAATTT-----ATCCACAGAAAGGAATATTTAAATCA 2969
DB CTCAGGTCCTCGCATACCAGAGCCTGGGGAATCCAGAGGGAGACATGTGAAGGCTCC 3414
QY 2970 CTGGATAAACAATATAAATAGCTTCATATTTGCTTCAATACCAGAACCATTT----- 3022
DB CTGGACCTCGTTCAGGACNGACCATGTCTGTGCCACGCGAGTCTCTAGGTACG 3474
QY 3023 -CAACTCTCTAGGTTTTAAAGTGGCTCGTCCGGAATTTGATCCCCCTCAGGATA----- 3074
DB CTAAATTTGGCCAGATCTACACTTGGTGGCAGCGAGTTTCAGGCTTCTGGGAAGCAAAAC 3534
QY 3075 -----TAGTAGTTTCGCTTTTGCATAGGAGGGGGAATGTAGTCTTATG 3119
DB TGGCAGAGAACAGACAGGATCTTGAAGTTGGGAAGGAAAGTCTAGGGCCAGAGACTGA 3594
QY 3120 CAATACTCTGTAGTCTTGAACATGTAACGATGAGTTAGCAATCCCTTACAAGGAG 3179
DB ACCTGGGCTGTCTGTCTTCCACTGTCTGCTCCGAGGTTTCATCTCCAGCTCTTACG 3654
QY 3180 AGAAAAAGCAGCGTCATCCGATTTGGTGAAGTAAGTGGTACGATCGTCCCTTATPAG 3239
DB CCCCTAGACCCAAAGCCAGCATGACGCCGCGAGTAGCAAGTAAAGAGCCACAGTCCAG 3714
QY 3240 GAAGGCA--CAGACGGCTGACATGGAATTTGAGAGAACCACTGAATTCGCAATGCGAG 3296
DB ACTGGGTAGGCGAGAGGTGCTGAGGCTTGGGCGAGGTGCTAGTTGGATCCGGTTTCCAG 3774
QY 3297 AGATATTGTATTA-----AGTCCCTAGCTCGATA 3326
DB AGGCATATATATATAAAGGTGCTGTTTCCCGATGTTGCAACACCCAGAGGCTTAGC 3834
QY 3327 CAGCAACGCCATTGACCATTTACCATTTGCACTGTTGCACTTCCAGCTTCACTGCGC 3386
DB CAGACCAACATTCCTGTCNAAAGCCAGCTCTCCATTTATAGCACTTTGGAAGACTAGCC 3894
QY 3387 CAAGCACTCAGGCGCAGAGGCTGCTAAAGAGGCGGAACAGGTA----- 3431
DB AAGGAGCTGAATGCAAGGCTTGAAGAGGACAGGGGGAATCAAGGGGCCACCTATGGC 3954
QY 3432 ----- 3431
DB GGCAGGAACTGAGCCTCAGATGAGCTGTGTCTTCCAGGTGAGTGGGCGCTACAGC 4014
QY 3432 -----GAAAGCCAGTCCGAGAAACGGTGTGACCCCGGATGAATGTGAGCTAC 3480
DB CTAAGAGCCCTTGCAATTCAGGGGCAATGCTCAGCCAGAGCCAAATGGCACGCTCCAGA 4074
QY 3481 TGGGCTATCTGCAAGGGAAGAAACGCAAGCGCAAGAGAAAGCAGG----- 3526
DB AGGGGTGGCTGGGGGAAGTTTGTAGTGACCATATAAATAAAGCAAGGTGCAATCACTTA 4134
QY 3527 -----TAGCTTGCAGTGGGCTT 3543
DB GACTCAGCATAAATTTATTTTCGGTTTTTNTAAATATTCTGTTTGTATTTGTGCTT 4194
QY 3544 ACATGGCGATAGCTAGACTGGGCGGTT----- 3571
DB GCTTTGCTTTGCTTGTGTTGCTGGNGTAGACACTGGAGAGGCTTGGGCGGCTGACCA 4254
QY 3572 -----TATGGA 3577
DB GATAGCCAAAGTGTACCGCGTATGGCTGCCAGTCTCTCTGCCCTCTCTCTAGTAGGCA 4314

QY 3578 CAGCAAGCGAAACCGAAATTCACAGTGGGCGCCCTCTGTGTAAGTTTGGGAAGCCCTGCA 3637
DB GCTGCCCCCAGCAGAGTTCGGGAGGGGGCACTCCCTTGGCCCCAGTGTACCTCGGG 4374
QY 3638 AAGTAAA-----CTGGATGGCTTTCTTGGCCGCCAAG 3668
DB ACCCAAGCTCCGCCCTACTACATCTCTATTTGGCTTGAGGGCGGGCTTTTAGGCTTACCCT 4434
QY 3669 GATCTGATGGCGCAGG-----GATCAAGATCTGATCAAGAGACAGGATG 3713
DB GGATAAATAGCCAGAGCTTGGGCTGCGAGTTCGAGGGCGGATCTGATCAAGAGACAGGATG 4494
QY 3714 AGGATCGTTTCGATGATTAAGATGGAATTCACGACGTTCTCCGGCGCTTGGGT 3773
DB AGGATCGTTTCGATGATTAAGATGGAATTCACGACGTTCTCCGGCGCTTGGGT 4554
QY 3774 GGAGAGGCTATTCCGCTATGACTGGGCAACAACAATTCGGCTGCTCTGATTCGCCCT 3833
DB GGAGAGGCTATTCCGCTATGACTGGGCAACAACAATTCGGCTGCTCTGATTCGCCCT 4614
QY 3834 GTTCCGGCTGTCAAGCAGGGGCGCCGCTTCTTTTGTCAAGACCGACCTGTCCGGTC 3893
DB GTTCCGGCTGTCAAGCAGGGGCGCCGCTTCTTTTGTCAAGACCGACCTGTCCGGTC 4674
QY 3894 CCTGAATGAACCTGACGACGAGGCGCGCTATCGTGGCTGGCCACGACGGGCTTCC 3953
DB CCTGAATGAACCTGACGACGAGGCGCGCTATCGTGGCTGGCCACGACGGGCTTCC 4734
QY 3954 TTGGCGAGCTGTGCTGACGTTGTCACTGAAGCGGGAAGGGAATGCGCTGCTATTGGGGCA 4013
DB TTGGCGAGCTGTGCTGACGTTGTCACTGAAGCGGGAAGGGAATGCGCTGCTATTGGGGCA 4794
QY 4014 AGTCCGGGCGAGGATCTCTGCTCATCTCACTTGTCTTCCCGCAGAAAGTATCCATCAT 4073
DB AGTCCGGGCGAGGATCTCTGCTCATCTCACTTGTCTTCCCGCAGAAAGTATCCATCAT 4854
QY 4074 GGCTGATGCAATGGGCGGCTGCATAGCTTTGATCCGGCTACCTGCCCATTCGACCACA 4133
DB GGCTGATGCAATGGGCGGCTGCATAGCTTTGATCCGGCTACCTGCCCATTCGACCACA 4914
QY 4134 AGCGAAACATCGCATTCGAGCGAGCAGCATCTCGATCGGAATGGAAGCCGCTTTGTCGATCAGGA 4193
DB AGCGAAACATCGCATTCGAGCGAGCAGCATCTCGATCGGAATGGAAGCCGCTTTGTCGATCAGGA 4974
QY 4194 TGATCTGCAAGAGAGCATCAGGGGCTCGGCCAGCGAACTGTTCCGACAGGCTCAAGGC 4253
DB TGATCTGCAAGAGAGCATCAGGGGCTCGGCCAGCGAACTGTTCCGACAGGCTCAAGGC 5034
QY 4254 GCGCATGCCGACCGCGAGGATCTCGTCTGACCCATCGCGATGCTGCTTGGCGAATAT 4313
DB GCGCATGCCGACCGCGAGGATCTCGTCTGACCCATCGCGATGCTGCTTGGCGAATAT 5094
QY 4314 CATGGTGAATAATGGCGCTTTTCTGATTCATCGACTGTGGCGGCTGGGTGGCGGGA 4373
DB CATGGTGAATAATGGCGCTTTTCTGATTCATCGACTGTGGCGGCTGGGTGGCGGGA 5154
QY 4374 CCCTATCAGGACATAGGTTGGCTACCCGTGATTTGCTGAAGGCTTGGCGGCGAATG 4433
DB CCCTATCAGGACATAGGTTGGCTACCCGTGATTTGCTGAAGGCTTGGCGGCGAATG 5214
QY 4434 GGCTGACCGCTTCTCTGCTTTTACGGTATCGCGCTCCCGATTCGACGCGCATCGCCTT 4493
DB GGCTGACCGCTTCTCTGCTTTTACGGTATCGCGCTCCCGATTCGACGCGCATCGCCTT 5274
QY 4494 CTATCGCTTCTGACGAGTTCTTCTGAGCGGGAATGAGAGCCCACTGTAGTTTGGCAAGAT 4553
DB CTATCGCTTCTTTCAGCAGTTCTTCTGAGCGGGAATGAGAGCCCACTGTAGTTTGGCAAGAT 5334
QY 4554 TTTATTTAGTCTCCAGAAAGGGGGGAATGAGAGCCCACTGTAGTTTGGCAAGCT 4613
DB TTTATTTAGTCTCCAGAAAGGGGGGAATGAGAGCCCACTGTAGTTTGGCAAGCT 5394
QY 4614 AGCTTAAGTAACGCCATTTTTCAGAGGATGGAATAATACATAACTGAGAAATAGAGAGTT 4673

Db 5395 AGCTTAAGTAACGCCATTTTGCNAGCATGGAAATACATACATGAGATAGAGAGTT 5454
Qy 4674 CAGATCAAGAGTTCAGGAACAGATGGAACAGCTGAATATGGGCCAAACAGGATATCTGTGGT 4733
Db 5455 CAGATCAAGTTCAGGAACAGATGGAACAGCTGAATATGGGCCAAACAGGATATCTGTGGT 5514
Qy 4734 AAGCAGTTCCTGCCCGGCTCAGGCCAAGAACAGATGGAACAGCTGAATATGGGCCAAA 4793
Db 5515 AAGCAGTTCCTGCCCGGCTCAGGCCAAGAACAGATGGAACAGCTGAATATGGGCCAAA 5574
Qy 4794 CAGGATATCTGTGGTAAGCAGTTCTGCCCGGCTCAGGCCAAGAACAGATGTCGCCA 4853
Db 5575 CAGGATATCTGTGGTAAGCAGTTCTGCCCGGCTCAGGCCAAGAACAGATGTCGCCA 5634
Qy 4854 GATCGGTTCAGCCCTCAGCAGTTCCTAGAGAACCATCAGATGTTTCCAGGGTTCGCCAA 4913
Db 5635 GATCGGTTCAGCCCTCAGCAGTTCCTAGAGAACCATCAGATGTTTCCAGGGTTCGCCAA 5694
Qy 4914 GGACCTGAATGAACCTGTGCTTATTTGAACCTAACCAATCAGTTTCTCGCTTCG 4973
Db 5695 GGACCTGAATGAACCTGTGCTTATTTGAACCTAACCAATCAGTTTCTCGCTTCG 5754
Qy 4974 TTGGCGGCTTCTGCTCCCGGCTCAATAAAGAGCCCAACCCCTCAGTTCGGGCGC 5033
Db 5755 TTGGCGGCTTCTGCTCCCGGCTCAATAAAGAGCCCAACCCCTCAGTTCGGGCGC 5814
Qy 5034 CAGTCTCCGATTCAGTCCCGGCTCAATAAAGAGCCCAACCCCTCAGTTCGGGCGC 5093
Db 5815 CAGTCTCCGATTCAGTCCCGGCTCAATAAAGAGCCCAACCCCTCAGTTCGGGCGC 5874
Qy 5094 TGCATCCAGTTCGCTGCTTCTTGGAGGGTCTCCTCGATGATGACTACC 5153
Db 5875 TGCATCCAGTTCGCTGCTTCTTGGAGGGTCTCCTCGATGATGACTACC 5934
Qy 5154 CGTACGGGGCTCTTCATTGG 5177
Db 5935 CGTACGGGGCTCTTCATTGG 5958

RESULT 3

US-09-380-190A-29
; Sequence 29, Application US/09380190A
; Patent No. 6410220
; GENERAL INFORMATION:
; APPLICANT: NATURE TECHNOLOGY CORPORATION, ET AL.
; TITLE OF INVENTION: SELF-ASSEMBLING GENES, VECTORS AND USES
; THEREOF
; NUMBER OF SEQUENCES: 74
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MUETING, RAASCH & GEBHARDT, P.A.
; STREET: 119 NORTH FOURTH STREET, SUITE 203
; CITY: MINNEAPOLIS
; STATE: MINNESOTA
; COUNTRY: USA
; ZIP: 55401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/380,190A
; FILING DATE: 26-Aug-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US98/03918
; FILING DATE: 28-FEB-98
; ATTORNEY/AGENT INFORMATION:
; NAME: MUETING, ANN M.
; REGISTRATION NUMBER: 33,977
; REFERENCE/DOCKET NUMBER: 228, 00010201
; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 612-305-1217
; TELEFAX: 612-305-1228
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5594 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 29:
US-09-380-190A-29

Query Match 31.5%; Score 1629.4; DB 3; Length 5594;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 1654; Conservative 0; Mismatches 21; Indels 2; Gaps 1;
Qy 3371 AAGCTTCAAGCTCCGCAAGCAGCTCAGGGCCAAAGGGCTGCTAAAGGAAGCGGAACAGT 3430
Db 3342 AAGCTTCAAGCTCCGCAAGCAGCTCAGGGCCAAAGGGCTGCTAAAGGAAGCGGAACAGT 3401
Qy 3431 AGAAAGCCAGTCCGCAAGAACCGGTGCTGACCCCGGATGAATGTGAGCTTACCTGGGCTATCT 3490
Db 3402 AGAAAGCCAGTCCGCAAGAACCGGTGCTGACCCCGGATGAATGTGAGCTTACCTGGGCTATCT 3461
Qy 3491 GGACAAGGGAAAACGCAAGCGCAAGAAAGAGAGAGTAGCTTGCAGTGGGGCTTACATGCC 3550
Db 3462 GGACAAGGGAAAACGCAAGCGCAAGAAAGAGAGAGTAGCTTGCAGTGGGGCTTACATGCC 3521
Qy 3551 GATAGCTAGACTGGGGCGGTTTTATGGACAGCAAGCAACCGGAATTGCGAGCTGGGGCGC 3610
Db 3522 GATAGCTAGACTGGGGCGGTTTTATGGACAGCAAGCAACCGGAATTGCGAGCTGGGGCGC 3581
Qy 3611 CCTCTGTTAAGTTGGGAAGCCCTGCAAGTAATAACTGGATGGCTTTCTTGGCCCAAGGA 3670
Db 3582 CCTCTGTTAAGTTGGGAAGCCCTGCAAGTAATAACTGGATGGCTTTCTTGGCCCAAGGA 3641
Qy 3671 TCTGATCGCGCAGGGGATCAAGATCTGATCAAGAGACAGATGAGATCGTTTGGCATGA 3730
Db 3642 TCTGATCGCGCAGGGGATCAAGATCTGATCAAGAGACAGATGAGATCGTTTGGCATGA 3701
Qy 3731 TTGAACAAGATGGATTGCAGCAGGTTCTCCGGCGGCTTGGGTGGAGAGGCTATTCCGCT 3790
Db 3702 TTGAACAAGATGGATTGCAGCAGGTTCTCCGGCGGCTTGGGTGGAGAGGCTATTCCGCT 3761
Qy 3791 ATGACTGGGCACACAGACAAATCCGGCTCTGATGCCCGCTGTTCCGGCTGTCCAGCGC 3850
Db 3762 ATGACTGGGCACACAGACAAATCCGGCTCTGATGCCCGCTGTTCCGGCTGTCCAGCGC 3821
Qy 3851 AGGGCGCCCGGTTCTTTTGTCAAGACCGACCTGTCCGGTGCCTGTAATGAATGCACTGCAGG 3910
Db 3822 AGGGCGCCCGGTTCTTTTGTCAAGACCGACCTGTCCGGTGCCTGTAATGAATGCACTGCAGG 3881
Qy 3911 ACGAGGACGCGCGCTATCGTGGCTGGCCAGACGCGGGGTTCTTGGCAGCTGTGCTCG 3970
Db 3882 ACGAGGACGCGCGCTATCGTGGCTGGCCAGACGCGGGGTTCTTGGCAGCTGTGCTCG 3941
Qy 3971 ACGTTGTCACTGAAGCGGAGGACCTGGCTGCTATTGGGGCGAAGTCCCGGGCAGGATC 4030
Db 3942 ACGTTGTCACTGAAGCGGAGGACCTGGCTGCTATTGGGGCGAAGTCCCGGGCAGGATC 4001
Qy 4031 TCTGTGATCTCACCTTGTCTCTCCCGGAGAAAGATTCATCATGCTGATGCAATGCGCGC 4090
Db 4002 TCTGTGATCTCACCTTGTCTCTCCCGGAGAAAGATTCATCATGCTGATGCAATGCGCGC 4061
Qy 4091 GGCTGCATACGCTTGATCCGGCTACCTGCCCATTTCCGACCACCAAGCGGAAACATCGCATCG 4150
Db 4062 GGCTGCATACGCTTGATCCGGCTACCTGCCCATTTCCGACCACCAAGCGGAAACATCGCATCG 4121
Qy 4151 AGCGAGCAGCTGACTCGGATGGAGCCGCTCTTGTGATCAGATGATCTGACCAAGAGC 4210
Db 4122 AGCGAGCAGCTGACTCGGATGGAGCCGCTCTTGTGATCAGATGATCTGACCAAGAGC 4181
Qy 4211 ATCAGGGGCTCGCGCCAGCCGAACTGTTCCGCGAGGCTCAAGGCGCGCATGCCCGAGCGC 4270

```
Db 4182 ATCAGGGGCTCGCGCAGCCGAACTGTTGCCAGGCTCAAGCGCGCATGCCGAGCGCG 4241
Qy 4271 AGGATCTCGTGTGACCCATGGCGATGCTGCTGCGCAATATATATGTTGAAATGGCC 4330
Db 4242 AGGATCTCGTGTGACCCATGGCGATGCTGCTGCGCAATATATATGTTGAAATGGCC 4301
Qy 4331 GCTTTCTCGGATTCATCCAGCTGTGCGCGCTGGGTGTGCGGACCGCTATCAGGACATAG 4390
Db 4302 GCTTTCTCGGATTCATCCAGCTGTGCGCGCTGGGTGTGCGGACCGCTATCAGGACATAG 4361
Qy 4391 CGTTGGCTACCCGTGATATGCTGAAGAGCTTGGCGGCAATGGGCTGACCGCTTCTCTCG 4450
Db 4362 CGTTGGCTACCCGTGATATGCTGAAGAGCTTGGCGGCAATGGGCTGACCGCTTCTCTCG 4421
Qy 4451 TGCTTTACGGTATCGCCGCTCCGATTCGACGCGCATCGCTTCTATGCGCTTCTTGACG 4510
Db 4422 TGCTTTACGGTATCGCCGCTCCGATTCGACGCGCATCGCTTCTATGCGCTTCTTGACG 4481
Qy 4511 AGTTCTTCTGAGCGGACTCTGGGGTTCGATAAATAAAGATTTTA--TTTAGTCTCCA 4568
Db 4482 AGTTCTTCTGAGCGGACTCTGGGGTTCGATAAATAAAGATTTTA--TTTAGTCTCCA 4541
Qy 4569 GAAAAAGGGGGAATGAAAGACCCACCTGTAGGTTTGGCAAGCTAGCTTAAAGTAAAGCC 4628
Db 4542 GAAAAAGGGGGAATGAAAGACCCACCTGTAGGTTTGGCAAGCTAGCTTAAAGTAAAGCC 4601
Qy 4629 ATTTTGAAGGATGGAATAATACATTAAGTATGAGAGTTCAGATCAAGGTGACG 4688
Db 4602 ATTTTGAAGGATGGAATAATACATTAAGTATGAGAGTTCAGATCAAGGTGACG 4661
Qy 4689 AACAGATGGAACAGCTGAAATGAGGCAACAGGATATCTGCTGTAAGCAGTCTCTGCC 4748
Db 4662 AACAGATGGAACAGCTGAAATGAGGCAACAGGATATCTGCTGTAAGCAGTCTCTGCC 4721
Qy 4749 CGGCTCAGGCGCAAGAACAGATGGAACAGCTGAAATGAGGCAACAGGATATCTGCTG 4808
Db 4722 CGGCTCAGGCGCAAGAACAGATGGAACAGCTGAAATGAGGCAACAGGATATCTGCTG 4781
Qy 4809 AAGCAGTTCTTCCCGGCTCAGGGCCAAAGATGGTCCCAGATGCGGTCCAGGCC 4868
Db 4782 AAGCAGTTCTTCCCGGCTCAGGGCCAAAGATGGTCCCAGATGCGGTCCAGGCC 4841
Qy 4869 TCAGCAGTTCTAGAGACCACTAGATGTTTCCAGGGTCCCAAGGACCTGAATGACC 4928
Db 4842 TCAGCAGTTCTAGAGACCACTAGATGTTTCCAGGGTCCCAAGGACCTGAATGACC 4901
Qy 4929 CTGTGCTTATTTGAATCAATCAATCAGTTCTGCTTCTGCTTCTGCTGCTTCTGCTG 4988
Db 4902 CTGTGCTTATTTGAATCAATCAATCAGTTCTGCTTCTGCTTCTGCTGCTTCTGCTG 4961
Qy 4989 TCCCGAGCTCAATAAAGAGCCCAACCCCTCACTCGGGGCGCAGTCTCCGAT 5045
Db 4962 TCCCGAGCTCAATAAAGAGCCCAACCCCTCACTCGGGGCGCAGTATCTGCT 5018
```

RESULT 4

```
US-09-897-511A-9
; Sequence 9, Application US/09897511A
; Patent No. 6852510
; GENERAL INFORMATION:
; APPLICANT: Bremel, Robert
; APPLICANT: Miller, Linda
; APPLICANT: Bleck, Gregory
; TITLE OF INVENTION: Host Cells Containing Multiple Integrating Vectors
; FILE REFERENCE: GALA-06416
; CURRENT APPLICATION NUMBER: US/09/897,511A
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 60/215,925
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
```

```
; LENGTH: 5130
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-09-897-511A-9

Query Match      30.7%; Score 1588.6; DB 3; Length 5130;
Best Local Similarity 66.1%; Pred. No. 0;
Matches 3007; Conservative 0; Mismatches 754; Indels 786; Gaps 18;

Qy 1 TTTGAAAGAGCCCAACCCGCTAGGTGGCAAGCTAGCTTAAAGTAAACGCCACTTTTGCAGGCAAT 60
Db 1 TTTGAAAGAGCCCAACCCGCTAGGTGGCAAGCTAGCTTAAAGTAAACGCCACTTTTGCAGGCAAT 60
Qy 61 GGAATAATACATTAATCAGAGATAGAAAAGTTTCAGATCAAGGTTCAGGAAACAAGAAACAGC 120
Db 61 GGAATAATACATTAATCAGAGATAGAAAAGTTTCAGATCAAGGTTCAGGAAACAAGAAACAGC 120
Qy 121 TGAATACCAACAGGATATCTGTGTAAGCGGTTTCTGCCCCCGCTCAGGGCCCAAGAAC 180
Db 121 TGAATACCAACAGGATATCTGTGTAAGCGGTTTCTGCCCCCGCTCAGGGCCCAAGAAC 180
Qy 181 GATGAGACAGCTGAGTGTGATGCGGCCAAACAGGATATCTGTGTAAGCAGTTCCTGCCCGG 240
Db 181 GATGAGACAGCTGAGTGTGATGCGGCCAAACAGGATATCTGTGTAAGCAGTTCCTGCCCGG 240
Qy 241 CTGCGGGCCAAAGAACAGATGCTGCCAGATGCGGTCCAGCCCTCAGCAGTTTCTAGTGAA 300
Db 241 CTGCGGGCCAAAGAACAGATGCTGCCAGATGCGGTCCAGCCCTCAGCAGTTTCTAGTGAA 300
Qy 301 TCATCAGATGTTTCCAGGGTCCCAAGGACCTGAAATGACCCCTGATCTTATTTGAAAC 360
Db 301 TCATCAGATGTTTCCAGGGTCCCAAGGACCTGAAATGACCCCTGATCTTATTTGAAAC 360
Qy 361 TAACCAATCAGTTTCGCTTCTGCTTCTGCTGCGCGCTTCCGCTCTCCGAGCTCAATAAA 420
Db 361 TAACCAATCAGTTTCGCTTCTGCTTCTGCTGCGCGCTTCCGCTCTCCGAGCTCAATAAA 420
Qy 421 AGAGCCCAACCCCTCACTCGCGCGCCAGTCTTCCGATAGACTGCGTCCCGGGGTAC 480
Db 421 AGAGCCCAACCCCTCACTCGCGCGCCAGTCTTCCGATAGACTGCGTCCCGGGGTAC 480
Qy 481 CCGTATTTCCCAATAAAGCCTCTTGTGTTGATCCGAAATCGTGCTCTCGCTGTTCTCTTG 540
Db 481 CCGTATTTCCCAATAAAGCCTCTTGTGTTGATCCGAAATCGTGCTCTCGCTGTTCTCTTG 540
Qy 541 GGAGGGTCTCTCTCTGAGTGAATGACTACCCAGACGAGCGGGGCTTTTCAATTTGGGGGCTCGT 600
Db 541 GGAGGGTCTCTCTCTGAGTGAATGACTACCCAGACGAGCGGGGCTTTTCAATTTGGGGGCTCGT 600
Qy 601 CCGGGATTTGGAGACCCCTCGCCAGGGACCAACAGCCACCCAGGGAGGTAACTGGCC 660
Db 601 CCGGGATTTGGAGACCCCTCGCCAGGGACCAACAGCCACCCAGGGAGGTAACTGGCC 660
Qy 661 AGCAACTTATCTGTGCTGCTGCGGATTCCTAGTCTATGTTTGTATGTTATGCGGCTCGG 720
Db 661 AGCAACTTATCTGTGCTGCTGCGGATTCCTAGTCTATGTTTGTATGTTATGCGGCTCGG 720
Qy 721 TCTGTACTAGTTAGTAACTAGCTCTGTATCTGCGGACCCCGTGGTGAACCTGACGAGTT 780
Db 721 TCTGTACTAGTTAGTAACTAGCTCTGTATCTGCGGACCCCGTGGTGAACCTGACGAGTT 780
Qy 781 CTGAACACCCGGCGCAACCCCTGGGAGACGCTCCAGGGACCTTTGGGGGCGCTTTTGTGG 840
Db 781 CTGAACACCCGGCGCAACCCCTGGGAGACGCTCCAGGGACCTTTGGGGGCGCTTTTGTGG 840
Qy 841 CCGACCTGAGGAGGAGTGCATGTTGGAATCCAGCCCGCTCAGGATATGTTGTTCTGCT 900
Db 841 CCGACCTGAGGAGGAGTGCATGTTGGAATCCAGCCCGCTCAGGATATGTTGTTCTGCT 900
Qy 901 AGGAGACGAGAACCTAAAAACAGTTTCCCGCTCTCCGCTCTGAAATTTTTCGTTTGGAA 960
Db 901 AGGAGACGAGAACCTAAAAACAGTTTCCCGCTCTCCGCTCTGAAATTTTTCGTTTGGAA 960
```


QY 3121 AATACTCTTGTAGTCTTTCACACATGGTAACGATGAGTTAGCAACATGCTTTACAGGAGA 3180
DB |||||
QY 2340 AATACATCTTGTAGTCTTTCACACATGGTAACGATGAGTTAGCAACATGCTTTACAGGAGA 2399
DB |||||
QY 3181 GAAAGACACCGTGTGATCCGATTTGGTGGAGTAAAGGTGGTACGATCGTCTTATTAGG 3240
DB |||||
QY 2400 GAAAGACACCGTGTGATCCGATTTGGTGGAGTAAAGGTGGTACGATCGTCTTATTAGG 2459
DB |||||
QY 3241 AAGCAACAGACGGTCTGACATGATTTGGAGCAACCACTGAATTCGCCATTCAGAGAT 3300
DB |||||
QY 2460 AAGCAACAGACGGTCTGACATGATTTGGAGCAACCACTGAATTCGCCATTCAGAGAT 2519
DB |||||
QY 3301 -ATTGTAATTAAGTCCCTAGCTCGATACAGCAACCGCA--TTTGACCAATTCACACAT 3357
DB |||||
QY 2520 AATTGTATTAAGTCCCTAGCTCGATACAGCAACCGCA--TTTGACCAATTCACACAT 2579
DB |||||
QY 3358 GGTGTGACCT--CCAAAGCTTTCAGCTGCGCAAGCACTCAAGGGCGCAAGGGCTGCTAAA 3415
DB |||||
QY 2580 GGTGTGACCTTCCAAAGCTTTCAGCTGCGCAAGCACTCAAGGGCGCAAGGGCTGCTAAA 2639
DB |||||
QY 3416 GGAAGCGCAACAGTGAAGACGAGTCCGCAAGCAAGCGGTGCTGACCCCGGATGAATGTCA 3475
DB |||||
QY 2640 GGAAGCGCAACAGTGAAGACGAGTCCGCAAGCAAGCGGTGCTGACCCCGGATGAATGTCA 2699
DB |||||
QY 3476 GCTACTGGCTATCTGGCAAGGGGAAACGCAAGCGCAAGAGAAAGCAGGTAGCTTGCA 3535
DB |||||
QY 2700 GCTACTGGCTATCTGGCAAGGGGAAACGCAAGCGCAAGAGAAAGCAGGTAGCTTGCA 2759
DB |||||
QY 3536 GTGGCTTACATGGCGATAGCTAGTCTGGCGGGTTTATGACAGCAAGCGAAACCGGAAT 3595
DB |||||
QY 2760 GTGGCTTACATGGCGATAGCTAGTCTGGCGGGTTTATGACAGCAAGCGAAACCGGAAT 2819
DB |||||
QY 3596 TGCCAGCTGGGGCGCCCTCTGTAAGGTTGGAGAGCCCTGCAAGTAACATGGATGGCTT 3655
DB |||||
QY 2820 TGCCAGCTGGGGCGCCCTCTGTAAGGTTGGAGAGCCCTGCAAGTAACATGGATGGCTT 2879
DB |||||
QY 3656 TCTTCCGCCAAGGATCTGATGGCGCAGGGATCAAGATCTGATCAAGAGACAGATGAG 3715
DB |||||
QY 2880 TCTTCCGCCAAGGATCTGATGGCGCAGGGATCAAGATCTGATCAAGAGACAGATGAG 2939
DB |||||
QY 3716 GATCGTTTCCGATGATGAACAAGATGATTCACGACGAGTTCTCCGCCGCTTGGGTGG 3775
DB |||||
QY 2940 GATCGTTTCCGATGATGAACAAGATGATTCACGACGAGTTCTCCGCCGCTTGGGTGG 2999
DB |||||
QY 3776 AGAGCTATTCCGCTATGATCGGGCAACAAGACATCGGCTGCTGATCGCCGCTGT 3835
DB |||||
QY 3000 AGAGCTATTCCGCTATGATCGGGCAACAAGACATCGGCTGCTGATCGCCGCTGT 3059
DB |||||
QY 3836 TCCGCTCTGACGCGAGGGCGCCCGGTTCTTTTGTCAAGACGACCTGTCCGCTGCC 3895
DB |||||
QY 3080 TCCGCTCTGACGCGAGGGCGCCCGGTTCTTTTGTCAAGACGACCTGTCCGCTGCC 3119
DB |||||
QY 3896 TGAATGAATGACGAGGAGGCGCGGCTATCGTGGCTGGCCACGACGGCGGCTTCTT 3955
DB |||||
QY 3120 TGAATGAATGACGAGGAGGCGCGGCTATCGTGGCTGGCCACGACGGCGGCTTCTT 3179
DB |||||
QY 3956 GCGCAGCTGTGCTGACGCTGTGATCAAGCGGGAGGAGGAGTGGCTGCTATTGGGCGAAG 4015
DB |||||
QY 3180 GCGCAGCTGTGCTGACGCTGTGATCAAGCGGGAGGAGGAGTGGCTGCTATTGGGCGAAG 3239
DB |||||
QY 4016 TGCCGGGCGAGGATCTCTGTCATCTCACTTGTCTTCTCCGCGAGAAAGTATCCATATGG 4075
DB |||||
QY 3240 TGCCGGGCGAGGATCTCTGTCATCTCACTTGTCTTCTCCGCGAGAAAGTATCCATATGG 3299
DB |||||
QY 4076 CTGATGCAATGCGGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4135
DB |||||
QY 3300 CTGATGCAATGCGGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 3359
DB |||||
QY 4136 CGAAACATCGCATCGAGCGAGCACTGATCGGATGGAAGCGCGGCTTGTGATCAGGATG 4195
DB |||||
QY 3360 CGAAACATCGCATCGAGCGAGCACTGATCGGATGGAAGCGCGGCTTGTGATCAGGATG 3419
DB |||||
QY 4196 ATCTGGACGAAGAGCATCAGGGGCTCGCGCCAGCGGCAACTGTGTCGCCAGGCTCAAGGCGC 4255

DB 3420 ATCTGGACGAAGACATCAGGGCTCGCGCCAGCGCAACTCTTCCAGGCTCAAGGCGC 3479
QY 4256 GCATGCCCGACGGCGAGGATCTCGTCTGAGCCCATGCGGATGCTTGGCGGAATATCA 4315
DB 3480 GCATGCCCGACGGCGAGGATCTCGTCTGAGCCCATGCGGATGCTTGGCGGAATATCA 3539
QY 4316 TGGTGGAAATGGCCGCTTTTCTGATTCATCGACTGTGGCGGCTGGGTGTCGCGGACC 4375
DB 3540 TGGTGGAAATGGCCGCTTTTCTGATTCATCGACTGTGGCGGCTGGGTGTCGCGGACC 3599
QY 4376 GCTATCAGGACATAGCGTTGGCTACCGGTGATTTGCTGAAGAGCTTGGCGGCAATGGG 4435
DB 3600 GCTATCAGGACATAGCGTTGGCTACCGGTGATTTGCTGAAGAGCTTGGCGGCAATGGG 3659
QY 4436 CTGACCGCTTCTCGTCTGCTTTACGGTATCGCGCTCCCGATTCGAGCGGATCGCTTCT 4495
DB 3660 CTGACCGCTTCTCGTCTGCTTTACGGTATCGCGCTCCCGATTCGAGCGCATCGCTTCT 3719
QY 4496 ATCGCTTCTTGACGAGTCTTCTTGAGCGGAGCTCTGGGGTTGATA 4542
DB 3720 ATCGCTTCTTGACGAGTCTTCTTGAGCGGAGCTCTGGGGTTGATA 3766

RESULT 5
US-08-352-990-1
; Sequence 1, Application US/08352990
; Patent No. 5681562
; GENERAL INFORMATION:
; APPLICANT: SOBOLE, ROBERT E
; APPLICANT: GAGE, FRED H
; APPLICANT: ROYSTON, IVOR
; APPLICANT: FRIEDMAN, THEODORE
; TITLE OF INVENTION: LYMPHOKINE GENE THERAPY OF CANCER
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CAMPBELL AND FLORES
; STREET: 4370 LA JOLLA VILLAGE DRIVE, SUITE 700
; CITY: SAN DIEGO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 91211
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/352,990
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/863,641
; FILING DATE: 19920403
; ATTORNEY/AGENT INFORMATION:
; NAME: CAMPBELL, CATHRYN
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-SD 9295
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-535-9001
; TELEFAX: 619-535-8949
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6365 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: complement (1..6365)

OTHER INFORMATION: /note= "Complementary strand of
FEATURE: PLXSN-RI-1L2"
NAME/KEY: CDS
LOCATION: 2557..3351
US-08-352-990-1

Query Match 28.7%; Score 1486; DB 2; Length 6365;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

| | | | |
|----|------|--|------|
| Qy | 3692 | GATCTGATCAAGAGACAGGATGAGATCGTTTCGATGATTAAGCAAGATGATGACG | 3751 |
| Db | 2522 | GATCTGATCAAGAGACAGGATGAGATCGTTTCGATGATTAAGCAAGATGATGACG | 2581 |
| Qy | 3752 | CAGGTTCTCCGGCCGCTTGGGTGGAGAGGCTATTGGCTATGACTGGGCACAAACAGAA | 3811 |
| Db | 2582 | CAGGTTCTCCGGCCGCTTGGGTGGAGAGGCTATTGGCTATGACTGGGCACAAACAGAA | 2641 |
| Qy | 3812 | TCGGCTGCTGATGCGCCGCTGTTCCGGCTGTGACGCGAGGGGCGCCGGTCTTTTGG | 3871 |
| Db | 2642 | TCGGCTGCTGATGCGCCGCTGTTCCGGCTGTGACGCGAGGGGCGCCGGTCTTTTGG | 2701 |
| Qy | 3872 | TCAGACCGACTGTCCGGTCCCTGATGAATGACTGCAGAGGAGCGCGGCTATCGT | 3931 |
| Db | 2702 | TCAGACCGACTGTCCGGTCCCTGATGAATGACTGCAGAGGAGCGCGGCTATCGT | 2761 |
| Qy | 3932 | GGCTGGCCACAGCGGGCGCTTCTTGGCAGAGCTGTGCTGATGCTCACTGAAGCGGAA | 3991 |
| Db | 2762 | GGCTGGCCACAGCGGGCGCTTCTTGGCAGAGCTGTGCTGATGCTCACTGAAGCGGAA | 2821 |
| Qy | 3992 | GGGACTGGCTGATTTGGGGGAGAGTGGCGGGGAGAGTCTCTGTCTATCATCTTGTCTC | 4051 |
| Db | 2822 | GGGACTGGCTGATTTGGGGGAGAGTGGCGGGGAGAGTCTCTGTCTATCATCTTGTCTC | 2881 |
| Qy | 4052 | CTGCGGAGAAAGTATCCATCATGCTGATGATCAATGGCGGGCTGCATACGCTTGCATCGG | 4111 |
| Db | 2882 | CTGCGGAGAAAGTATCCATCATGCTGATGATCAATGGCGGGCTGCATACGCTTGCATCGG | 2941 |
| Qy | 4112 | CTACTCGCCATTTCGACCAACCAAGCGAAATCGCATCGAGCGAGCATCTCGGATGG | 4171 |
| Db | 2942 | CTACTCGCCATTTCGACCAACCAAGCGAAATCGCATCGAGCGAGCATCTCGGATGG | 3001 |
| Qy | 4172 | AAGCGGCTTTTCGATCAGGATGATCTGGACGAGAGCATCAGGGGCTCGCGCAGCG | 4231 |
| Db | 3002 | AAGCGGCTTTTCGATCAGGATGATCTGGACGAGAGCATCAGGGGCTCGCGCAGCG | 3061 |
| Qy | 4232 | AAGTGTTCGACAGGCTCAAGGGCGCATGCGCGAGGAGATCTGCTGTGACCCATG | 4291 |
| Db | 3062 | AAGTGTTCGACAGGCTCAAGGGCGCATGCGCGAGGAGATCTGCTGTGACCCATG | 3121 |
| Qy | 4292 | GCGATGCTGCTTTCGGAATATCATGGTGGAAATGCGCGCTTTCTGGATTCATCGACT | 4351 |
| Db | 3122 | GCGATGCTGCTTTCGGAATATCATGGTGGAAATGCGCGCTTTCTGGATTCATCGACT | 3181 |
| Qy | 4352 | GTGGCGGCTGGGTGTGCGGACCGCTATCAGGACATAGCGTTGCTACCCGCTGATTTG | 4411 |
| Db | 3182 | GTGGCGGCTGGGTGTGCGGACCGCTATCAGGACATAGCGTTGCTACCCGCTGATTTG | 3241 |
| Qy | 4412 | CTGAAGAGCTTGGCGGGAATGGGCTGACCGCTTCTGCTGCTTTACGGTATCGCGCTC | 4471 |
| Db | 3242 | CTGAAGAGCTTGGCGGGAATGGGCTGACCGCTTCTGCTGCTTTACGGTATCGCGCTC | 3301 |
| Qy | 4472 | CCGATTCGACGGATCGCTTCTATCGCTTCTGAGAGTCTTCTGAGCGGACTCT | 4531 |
| Db | 3302 | CCGATTCGACGGATCGCTTCTATCGCTTCTGAGAGTCTTCTGAGCGGACTCT | 3361 |
| Qy | 4532 | GGGGTTCGATAAAATAAAGATTTTATTAGTCTCCAGAAAAGGGGGAATGAAGACC | 4591 |
| Db | 3362 | GGGGTTCGATAAAATAAAGATTTTATTAGTCTCCAGAAAAGGGGGAATGAAGACC | 3421 |
| Qy | 4592 | CCACTGTAGGTTGGCAGCTAGCTTAAGTAACGCCATTTTTCAGAGCATGGAATAA | 4651 |

| | | | |
|----|------|--|------|
| Db | 3422 | CCACTGTAGGTTGGCAAGCTAGCTTAAGTAACGCCATTTTTCAGAGCATGGAATAA | 3481 |
| Qy | 4652 | CATTAAGTGAATAGAGAGGTTTCAAGTCAAGGTAAGGTAAGTGAAGTGAATATG | 4711 |
| Db | 3482 | CATTAAGTGAATAGAGAGGTTTCAAGTCAAGGTAAGGTAAGTGAAGTGAATATG | 3541 |
| Qy | 4712 | GGCCAAACAGGATATCTGTGTAGCAGTTCCTCCCGGCTCAGGGCCAAAGAACAGATG | 4771 |
| Db | 3542 | GGCCAAACAGGATATCTGTGTAGCAGTTCCTCCCGGCTCAGGGCCAAAGAACAGATG | 3601 |
| Qy | 4772 | GAAACAGCTGAATATGGGCCAAACAGGATATCTGTGTAAAGCAGTTCCTCCCGGCTCAG | 4831 |
| Db | 3602 | GAAACAGCTGAATATGGGCCAAACAGGATATCTGTGTAAAGCAGTTCCTCCCGGCTCAG | 3661 |
| Qy | 4832 | GGCCAAACAGATGCTGCCAGATGCGGTCACGCCCTCAGCAGTTCCTAGAGAACCATC | 4891 |
| Db | 3662 | GGCCAAACAGATGCTGCCAGATGCGGTCACGCCCTCAGCAGTTCCTAGAGAACCATC | 3721 |
| Qy | 4892 | AGATGTTTCCAGGGTCCCAAGGACCTGAAATGACCTGTGCTGCTTATTTGAACCTAACCA | 4951 |
| Db | 3722 | AGATGTTTCCAGGGTCCCAAGGACCTGAAATGACCTGTGCTGCTTATTTGAACCTAACCA | 3781 |
| Qy | 4952 | ATCAGTTCGCTTCTCGCTTCTGTTGCGCGCTTCTGCTCCCGAGCTCAATAAAGAGGCC | 5011 |
| Db | 3782 | ATCAGTTCGCTTCTCGCTTCTGTTGCGCGCTTCTGCTCCCGAGCTCAATAAAGAGGCC | 3841 |
| Qy | 5012 | CACAAACCCCTCACTCGGGGCGCCAGTCTCCGATGATGATCGCGGCTACCCGGT | 5071 |
| Db | 3842 | CACAAACCCCTCACTCGGGGCGCCAGTCTCCGATGATGATCGCGGCTACCCGGT | 3901 |
| Qy | 5072 | ATCAATAAACCCTTTCGAGTTCGATCCGATCCGATTCGCTGCTGCTGCTTCTTGGAGGGT | 5131 |
| Db | 3902 | ATCAATAAACCCTTTCGAGTTCGATCCGATCCGATTCGCTGCTGCTGCTTCTTGGAGGGT | 3961 |
| Qy | 5132 | CTCCTCTGAGTGTGATGACTACCCGCTCAGCGGGGCTCTTTTCAATTTGG | 5177 |
| Db | 3962 | CTCCTCTGAGTGTGATGACTACCCGCTCAGCGGGGCTCTTTTCAATTTGG | 4007 |

RESULT 6

US-08-786-531B-3

; Sequence 3, Application US/08786531B

; Patent No. 6541197

; GENERAL INFORMATION:

; APPLICANT: Link, Charles J.

; APPLICANT: Levy, John P.

; APPLICANT: Wang, Suming

; APPLICANT: Seragina, Tatiana

; TITLE OF INVENTION: Vehicles for Stable Transfer of Green

; NUMBER OF SEQUENCES: 10

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Zarley, McKee, Thomte, Voorhees & Sease

; STREET: 801 Grand Suite 3200

; CITY: Des Moines

; STATE: Iowa

; COUNTRY: United States

; ZIP: 50309

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA: US/08/786,531B

; FILING DATE: 21-JAN-1997

; CLASSIFICATION: 435

; PRIOR APPLICATION NUMBER: US 60/010371

; FILING DATE: 22-JAN-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: Nebel, Heidi S.

; REGISTRATION NUMBER: 37,719

REFERENCE/DOCKET NUMBER: hqtri
TELECOMMUNICATION INFORMATION:
TELEPHONE: 515-288-3667
TELEFAX: 515-288-1338
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 6620 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-786-531B-3

Query Match 28.7%; Score 1486; DB 3; Length 6620;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

| | | | |
|----|------|---|------|
| QY | 3692 | GATCTGATCAAGACAGCAGGATGAGGATCGTTTCGCATGATTGACACAGATGGATTGCAAG | 3751 |
| DB | 2777 | GATCTGATCAAGACAGCAGGATGAGGATCGTTTCGCATGATTGACACAGATGGATTGCAAG | 2836 |
| QY | 3752 | CAGGTTCTCCCGCGCTTGGGTGAGAGGCTATTCCGCTATGACTGGGCAACAACAGACAA | 3811 |
| DB | 2837 | CAGGTTCTCCCGCGCTTGGGTGAGAGGCTATTCCGCTATGACTGGGCAACAACAGACAA | 2896 |
| QY | 3812 | TCGGCTGCTGATGTCGCCCGCTGTTCCGGCTGTTCAGCGCAGGGCGCCGGTCTTTTTCG | 3871 |
| DB | 2897 | TCGGCTGCTGATGTCGCCCGCTGTTCCGGCTGTTCAGCGCAGGGCGCCGGTCTTTTTCG | 2956 |
| QY | 3872 | TCAGACCGCCTGTCGGGTGCCCTGATGATGAATGCAAGGACGAGCGCGGCTATCGT | 3931 |
| DB | 2957 | TCAGACCGCCTGTCGGGTGCCCTGATGATGAATGCAAGGACGAGCGCGGCTATCGT | 3016 |
| QY | 3932 | GGCTGGCCACAGCGGGCTTCTTGGCGCAGCTGTCTGACGTTGTCACTGAAGCGGAA | 3991 |
| DB | 3017 | GGCTGGCCACAGCGGGCTTCTTGGCGCAGCTGTCTGACGTTGTCACTGAAGCGGAA | 3076 |
| QY | 3992 | GGGACTGGCTGCTATTGGGCGAAGTGGCGGGCAGGATCTCTGTCATCTCACCTTGCTC | 4051 |
| DB | 3077 | GGGACTGGCTGCTATTGGGCGAAGTGGCGGGCAGGATCTCTGTCATCTCACCTTGCTC | 3136 |
| QY | 4052 | CTGCCGAGAAATGATCATGCTGATGCTGATGCAATGCGCGGCTGATACGCTTGATCGG | 4111 |
| DB | 3137 | CTGCCGAGAAATGATCATGCTGATGCTGATGCAATGCGCGGCTGATACGCTTGATCGG | 3196 |
| QY | 4112 | CTACTCGCCATTGACACCAAGCGAAATCATGCTGACGAGCAGCTACTCGGATGG | 4171 |
| DB | 3197 | CTACTCGCCATTGACACCAAGCGAAATCATGCTGACGAGCAGCTACTCGGATGG | 3256 |
| QY | 4172 | AAGCCGGTCTTTGCGATCAGGATGATCTGGACGAGCAGCATCAGGGGCTCGCCAGCCG | 4231 |
| DB | 3257 | AAGCCGGTCTTTGCGATCAGGATGATCTGGACGAGCAGCATCAGGGGCTCGCCAGCCG | 3316 |
| QY | 4232 | AAGTGTTCGCCAGGCTCAAGGGCGCATGCGCGAGCGGAGGATCTGCTGTCGACCCATG | 4291 |
| DB | 3317 | AAGTGTTCGCCAGGCTCAAGGGCGCATGCGCGAGCGGAGGATCTGCTGTCGACCCATG | 3376 |
| QY | 4292 | GCGATGCTGCTTGGCGGAATCATGCTGGAATGAGCGGCTTTTCTGGAATTCATCGACT | 4351 |
| DB | 3377 | GCGATGCTGCTTGGCGGAATCATGCTGGAATGAGCGGCTTTTCTGGAATTCATCGACT | 3436 |
| QY | 4352 | GTGGCCGCTGGGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCGCTGATATTG | 4411 |
| DB | 3437 | GTGGCCGCTGGGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCGCTGATATTG | 3496 |
| QY | 4412 | CTGAAGAGCTGGCGGCAATGGGCTGACCGCTTCTGCTGCTTTTACGGTATCGCCGCTC | 4471 |
| DB | 3497 | CTGAAGAGCTGGCGGCAATGGGCTGACCGCTTCTGCTGCTTTTACGGTATCGCCGCTC | 3556 |
| QY | 4472 | CCGATTCGAGCGGATCGCCCTTCTATCGCTTCTTTCAGAGTTCCTCTGAGCGGACTCT | 4531 |

| | | | |
|----|------|---|------|
| DB | 3557 | CCGATTCGACGCGCATCGCCTTCTTATCGCCTCTTGACGAGTTCTCTGACGGGACTCT | 3616 |
| QY | 4532 | GGGTTTCGATAAAATAAAAGATTTTATTTAGTCTCCAGAAAAGGGGGGATGAAGACC | 4591 |
| DB | 3617 | GGGTTTCGATAAAATAAAAGATTTTATTTAGTCTCCAGAAAAGGGGGGATGAAGACC | 3676 |
| QY | 4592 | CCACCTGTAGGTTTGGCAAGCTAGCTTAAAGTAAAGCCATTTTGAAGGCAATGAAAAATA | 4651 |
| DB | 3677 | CCACCTGTAGGTTTGGCAAGCTAGCTTAAAGTAAAGCCATTTTGAAGGCAATGAAAAATA | 3736 |
| QY | 4652 | CATAACTGAGAAATAGAGAAATTCAGATCAAGGTCAGAAACAGATGGAACAGCTGAATATG | 4711 |
| DB | 3737 | CATAACTGAGAAATAGAGAAATTCAGATCAAGGTCAGAAACAGATGGAACAGCTGAATATG | 3796 |
| QY | 4712 | GGCCAAACAGGATATCTGTGTAAGCAGTTTCTGCCCGGCTCAGGGCCAAAGAACAGATG | 4771 |
| DB | 3797 | GGCCAAACAGGATATCTGTGTAAGCAGTTTCTGCCCGGCTCAGGGCCAAAGAACAGATG | 3856 |
| QY | 4772 | GAACAGCTGAATATGGGCCAAACAGGATATCTGTGTAAGCAGTTTCTGCCCGGCTCAG | 4831 |
| DB | 3857 | GAACAGCTGAATATGGGCCAAACAGGATATCTGTGTAAGCAGTTTCTGCCCGGCTCAG | 3916 |
| QY | 4832 | GGCCAAAGACAGATGGTCCCGCAGATGCGTCCAGCCTCAGCAGTTTCTAGAGAACCATC | 4891 |
| DB | 3917 | GGCCAAAGACAGATGGTCCCGCAGATGCGTCCAGCCTCAGCAGTTTCTAGAGAACCATC | 3976 |
| QY | 4892 | AGATGTTTCCAGGGTGGCCCAAGGACCTGAATGACCTGTGCTTATTTTGAACCTAACCA | 4951 |
| DB | 3977 | AGATGTTTCCAGGGTGGCCCAAGGACCTGAATGACCTGTGCTTATTTTGAACCTAACCA | 4036 |
| QY | 4952 | ATCAGTTTCGGTTCCTCGGTTCTGTTGCGCGCTTCTGCTGCCGAGCTCAATAAAGAGCC | 5011 |
| DB | 4037 | ATCAGTTTCGGTTCCTCGGTTCTGTTGCGCGCTTCTGCTGCCGAGCTCAATAAAGAGCC | 4096 |
| QY | 5012 | CACAAACCCCTCACTCGGGGCGCCAGATCCTCCGATTTGACTGAGTCGCCCGGTACCCGCT | 5071 |
| DB | 4097 | CACAAACCCCTCACTCGGGGCGCCAGATCCTCCGATTTGACTGAGTCGCCCGGTACCCGCT | 4156 |
| QY | 5072 | ATCCAAATAAACCTCTTTGAGTTGCTGATCCGACTTGGTCTGCTGCTTCTTGGAGGGT | 5131 |
| DB | 4157 | ATCCAAATAAACCTCTTTGAGTTGCTGATCCGACTTGGTCTGCTGCTTCTTGGAGGGT | 4216 |
| QY | 5132 | CTCCTCTGAGTGTGATGACTACCGTACCGGCGGCTCTTTCATTTGG | 5177 |
| DB | 4217 | CTCCTCTGAGTGTGATGACTACCGTACCGGCGGCTCTTTCATTTGG | 4262 |

RESULT 7
US-07-753-520B-3
Sequence 3, Application US/07753520B
Patent No. 5352595
GENERAL INFORMATION:
APPLICANT: Tapscott, J.; Weintraub, H. M.; Palmer, T. D.
TITLE OF INVENTION: "MYOD REGULATORY REGION"
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Christensen, O'Connor, Johnson and Kindness
STREET: 2800 Pacific First Center, 1420 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98101-2347
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette-5.25 inch, 1.2Mb storage
COMPUTER: IBM PC/386 Compatible
OPERATING SYSTEM: MS-DOS 4.01
SOFTWARE: Word for Windows-t
CURRENT APPLICATION DATA: US/07753, 520B
APPLICATION NUMBER: 19910903
FILING DATE: 19910903
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: none

FILLING DATE: none
ATTORNEY/AGENT INFORMATION:
NAME: Sundemo, John, S.
REGISTRATION NUMBER: 34,446
REFERENCE/DOCKET NUMBER: PHCR-1-5789
TELECOMMUNICATION INFORMATION:
TELEPHONE: 1-206-682-8100; 1-206-224-0727 (direct)
TELEFAX: 1-206-224-0779
TELEX: 4938023
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 9115 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Other; plasmid DNA
DESCRIPTION: PLHMDM-53: 5'LTR (position 1-1159); Y+ (position 1159-1640); Hsd (position 1641-3007); Myo-D 531.4 Apat fragment (position 3008-5248); driving neo (position 5249-6117); with a 3'LTR (position 6118-6823) coupled to a pBR322 plasmid (position 6824-9115); Figures 7A-7D.
US-07-753-520B-3

Query Match 28.7%; Score 1486; DB 2; Length 9115;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

| | | | |
|----|------|--|------|
| Qy | 3692 | GATCTGATCAAGACAGACGATGAGATCGTTTCGATGATTGAACAAGATGGATTGCAAG | 3751 |
| Db | 5272 | GATCTGATCAAGACAGACGATGAGATCGTTTCGATGATTGAACAAGATGGATTGCAAG | 5331 |
| Qy | 3752 | CAGTTCTCCGGCGCTTGGTGGAGAGGCTATTTCGGCTATGATGGGCAACAAGACAA | 3811 |
| Db | 5332 | CAGTTCTCCGGCGCTTGGTGGAGAGGCTATTTCGGCTATGATGGGCAACAAGACAA | 5391 |
| Qy | 3812 | TCGGCTGCTGATGCGCGCTGTTCCGGCTGTCAGCGCAGGGCGCGCTTTCTTTTG | 3871 |
| Db | 5392 | TCGGCTGCTGATGCGCGCTGTTCCGGCTGTCAGCGCAGGGCGCGCTTTCTTTTG | 5451 |
| Qy | 3872 | TCAAGACCGACTGTCCGGTGCCTGATGAATGACGAGGAGCGGCGGCTATCGT | 3931 |
| Db | 5452 | TCAAGACCGACTGTCCGGTGCCTGATGAATGACGAGGAGCGGCGGCTATCGT | 5511 |
| Qy | 3932 | GGCTGGGCAACAAGCGGCTTCTTGGCGAGCTGTGCTCGAGTTGTCACTGAAGCGGAA | 3991 |
| Db | 5512 | GGCTGGGCAACAAGCGGCTTCTTGGCGAGCTGTGCTCGAGTTGTCACTGAAGCGGAA | 5571 |
| Qy | 3992 | GGGACTGGCTGATTGGGCGAAGTCCGGGCGAGGATCTCTCATCTCACCTTGGTC | 4051 |
| Db | 5572 | GGGACTGGCTGATTGGGCGAAGTCCGGGCGAGGATCTCTCATCTCACCTTGGTC | 5631 |
| Qy | 4052 | CTGCCGAGAAAGTATCCATCATGCTGATGCAATGCGCGGCTGCATACGCTTCATCCGG | 4111 |
| Db | 5632 | CTGCCGAGAAAGTATCCATCATGCTGATGCAATGCGCGGCTGCATACGCTTCATCCGG | 5691 |
| Qy | 4112 | CTACTCTGCCATTCGACCAACAGCGAAACATCGCATCGAGCGAGCAGTACTCGGATGG | 4171 |
| Db | 5692 | CTACTCTGCCATTCGACCAACAGCGAAACATCGCATCGAGCGAGCAGTACTCGGATGG | 5751 |
| Qy | 4172 | AAGCGGCTTGTGATCAGATGATCTGGACGAGAGCATCAGGGGCTCGCGCAGCGG | 4231 |
| Db | 5752 | AAGCGGCTTGTGATCAGATGATCTGGACGAGAGCATCAGGGGCTCGCGCAGCGG | 5811 |
| Qy | 4232 | AACGTGTCGAGGCTCAAGCGGCGCATGCGCGGCGAGGATCTCGTGTGACCCATG | 4291 |
| Db | 5812 | AACGTGTCGAGGCTCAAGCGGCGCATGCGCGGCGAGGATCTCGTGTGACCCATG | 5871 |
| Qy | 4292 | GCGATGCTGCTTCCGAAATATCATGTGGAAATGCGCGCTTTTCTGGATTCACTGACT | 4351 |
| Db | 5872 | GCGATGCTGCTTCCGAAATATCATGTGGAAATGCGCGCTTTTCTGGATTCACTGACT | 5931 |
| Qy | 4352 | GTGCGCGGCTGGGTGGCGGACCGCTATCAGGACATAGCGTTGCTACCGTGATTTG | 4411 |

| | | | |
|----|------|---|------|
| Db | 5932 | GTGCGCGGCTGGGTGGCGGACCGCTATCAGGACATAGCGTTGCTACCGTGATTTG | 5991 |
| Qy | 4412 | CTGAAGAGCTTGGCGGCAATGGGCTCAACCGCTTCTCGTGTCTTACGGTATCCCGCTC | 4471 |
| Db | 5992 | CTGAAGAGCTTGGCGGCAATGGGCTCAACCGCTTCTCGTGTCTTACGGTATCCCGCTC | 6051 |
| Qy | 4472 | CCGATTCGACGCGCATCGCTTCTATCGCTTCTTGAAGAGTCTTCTGAGCGGAGCTCT | 4531 |
| Db | 6052 | CCGATTCGACGCGCATCGCTTCTATCGCTTCTTGAAGAGTCTTCTGAGCGGAGCTCT | 6111 |
| Qy | 4532 | GGGTTTCGATAAAATAAAGATTTTATTAGTCTCCAGAAAAAGGGGGAATGAAAGACC | 4591 |
| Db | 6112 | GGGTTTCGATAAAATAAAGATTTTATTAGTCTCCAGAAAAAGGGGGAATGAAAGACC | 6171 |
| Qy | 4592 | CCACCTTAGCTTTGGCAAGCTAGCTTTAAGTAAAGCCATTTTGAAGGCAATGAAAAATA | 4651 |
| Db | 6172 | CCACCTTAGCTTTGGCAAGCTAGCTTTAAGTAAAGCCATTTTGAAGGCAATGAAAAATA | 6231 |
| Qy | 4652 | CATTAACCTGAGATAGAGAACTTCAAGATCAAGGTCAGGACAGATGGAACAGCTGAATG | 4711 |
| Db | 6232 | CATTAACCTGAGATAGAGAACTTCAAGATCAAGGTCAGGACAGATGGAACAGCTGAATG | 6291 |
| Qy | 4712 | GGCCAAACAGGATATCTGTGTAGCAGTTCCTCCCGCGCTCAGGGCCAAAGAACAGATG | 4771 |
| Db | 6292 | GGCCAAACAGGATATCTGTGTAGCAGTTCCTCCCGCGCTCAGGGCCAAAGAACAGATG | 6351 |
| Qy | 4772 | GAACAGCTGAATATGGGCCCAAAACAGGATATCTGTGTAGCAGTTCCTCCCGCGCTCAG | 4831 |
| Db | 6352 | GAACAGCTGAATATGGGCCCAAAACAGGATATCTGTGTAGCAGTTCCTCCCGCGCTCAG | 6411 |
| Qy | 4832 | GGCCAAACAGATGTTCCCGAGATGCGGTCAGGCTCAGGCTCAGGAGTTCCTAGAGAACATC | 4891 |
| Db | 6412 | GGCCAAACAGATGTTCCCGAGATGCGGTCAGGCTCAGGCTCAGGAGTTCCTAGAGAACATC | 6471 |
| Qy | 4892 | AGATGTTTCCAGGTCGCCCAAGGACCTGAAATGACCTGTCCTTATTTGAACTAAACA | 4951 |
| Db | 6472 | AGATGTTTCCAGGTCGCCCAAGGACCTGAAATGACCTGTCCTTATTTGAACTAAACA | 6531 |
| Qy | 4952 | ATCAGTTTCGCTTCTCGCTTCTGTTGCGCGGCTTCTGCTCCCGAGCTCAATAAAGAGCC | 5011 |
| Db | 6532 | ATCAGTTTCGCTTCTCGCTTCTGTTGCGCGGCTTCTGCTCCCGAGCTCAATAAAGAGCC | 6591 |
| Qy | 5012 | CACAAACCTTCACTCGGGGCGCAGTCTCCGATGATGATGTCGCGGCTACCGGTGT | 5071 |
| Db | 6592 | CACAAACCTTCACTCGGGGCGCAGTCTCCGATGATGATGTCGCGGCTACCGGTGT | 6651 |
| Qy | 5072 | ATCCAAATAAACCTTCTGAGTTGTCATCCGACTTGTGCTCTGCTGTTCTTGGGAGGT | 5131 |
| Db | 6652 | ATCCAAATAAACCTTCTGAGTTGTCATCCGACTTGTGCTCTGCTGTTCTTGGGAGGT | 6711 |
| Qy | 5132 | CTCCTCTGAGTGTGACTACCCGCTCAGCGGGGCTCTTTTCATTTGG | 5177 |
| Db | 6712 | CTCCTCTGAGTGTGACTACCCGCTCAGCGGGGCTCTTTTCATTTGG | 6757 |

RESULT 8

US-08-654-737B-3
; Sequence 3, Application US/08654737B
; Patent No. 6274136
; GENERAL INFORMATION:
; APPLICANT: University of Southern California
; APPLICANT: Weiner, Leslie P.
; APPLICANT: McMillan, Minnie
; TITLE OF INVENTION: CONSTRUCTION AND USE OF GENES ENCODING
; FILE REFERENCE: 13761-703-00 US
; CURRENT FILING DATE: 1996-05-29
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 5865
; TYPE: DNA

| | | | | | | | | | |
|--|------|---|------|--|--|--|--|--|--|
| ; ORGANISM: Artificial Sequence | | | | | | | | | |
| ; FEATURE: | | | | | | | | | |
| ; OTHER INFORMATION: Retroviral vector derived from Moloney Murine | | | | | | | | | |
| ; OTHER INFORMATION: Leukemia Virus | | | | | | | | | |
| US-08-654-737B-3 | | | | | | | | | |
| Query Match 28.1%; Score 1455.2; DB 3; Length 5865; | | | | | | | | | |
| Best Local Similarity 98.4%; Pred. No. 0; | | | | | | | | | |
| Matches 1488; Conservative 0; Mismatches 3; Indels 21; Gaps 1; | | | | | | | | | |
| Qy | 3687 | ATCAAGATCTGATCAAGACAGAGGATGAGGATCGTTTCGCATGATTTGAACAAAGATGGAATT | 3746 | | | | | | |
| Db | 1852 | ATTCGGATCTGATCAAGACAGAGGATGAGGATCGTTTCGCATGATTTGAACAAAGATGGAATT | 1911 | | | | | | |
| Qy | 3747 | GCAGCAGGTTCTCCGGCCGCTTGCGTGAGAGGCTATTCCGCTATGACTGGGCACACACA | 3806 | | | | | | |
| Db | 1912 | GCACGACGTTCTCCGGCCGCTTGCGTGAGAGGCTATTCCGCTATGACTGGGCACACACA | 1971 | | | | | | |
| Qy | 3807 | GACAACTGGCTGCTCTGATGCCCGCTGTTCCGGCTGTGAGCGAGGGCGCCCGGTTCT | 3866 | | | | | | |
| Db | 1972 | GACAACTGGCTGCTCTGATGCCCGCTGTTCCGGCTGTGAGCGAGGGCGCCCGGTTCT | 2031 | | | | | | |
| Qy | 3867 | TTTTGTCAAGACCGACTGTCGGTGCCTGAATGAATGACGACGAGGCGCGGCT | 3926 | | | | | | |
| Db | 2032 | TTTTGTCAAGACCGACTGTCGGTGCCTGAATGAATGACGACGAGGCGCGGCT | 2091 | | | | | | |
| Qy | 3927 | ATCGTGGCTGGCCAGCAGCGGGCTTCTTGCGAGCTGTGCTCGACGTTGTCACTGAAGC | 3986 | | | | | | |
| Db | 2092 | ATCGTGGCTGGCCAGCAGCGGGCTTCTTGCGAGCTGTGCTCGACGTTGTCACTGAAGC | 2151 | | | | | | |
| Qy | 3987 | GGGAAGGACTGGCTGCTATTGGGCGAAGTGC CGGGGAGGATCTCTGTCTCATCTCACCT | 4046 | | | | | | |
| Db | 2152 | GGGAAGGACTGGCTGCTATTGGGCGAAGTGC CGGGGAGGATCTCTGTCTCATCTCACCT | 2211 | | | | | | |
| Qy | 4047 | TGCTCTCCGAGAAAGTATCATATGCTGTGATGAATGCGCGCTGTCATACGCTTGA | 4106 | | | | | | |
| Db | 2212 | TGCTCTCCGAGAAAGTATCATATGCTGTGATGAATGCGCGCTGTCATACGCTTGA | 2271 | | | | | | |
| Qy | 4107 | TCCGGCTACCTGCCATTCGACCCAGCCAGCGAATCGCATCGAGCGAGCACCTACTCG | 4166 | | | | | | |
| Db | 2272 | TCCGGCTACCTGCCATTCGACCCAGCCAGCGAATCGCATCGAGCGAGCACCTACTCG | 2331 | | | | | | |
| Qy | 4167 | GATGAAGCCGGTCTTGTGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCGGGCC | 4226 | | | | | | |
| Db | 2332 | GATGAAGCCGGTCTTGTGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCGGGCC | 2391 | | | | | | |
| Qy | 4227 | AGCGGAATGTTGCGCAGGCTCAAGGCGCGCATGCCGACCGGAGGATCTCGTCGTGAC | 4286 | | | | | | |
| Db | 2392 | AGCGGAATGTTGCGCAGGCTCAAGGCGCGCATGCCGACCGGAGGATCTCGTCGTGAC | 2451 | | | | | | |
| Qy | 4287 | CCATGGCGATCGCTGTTGCCGAATATCATGGTGGAAATGGCCGCTTTTCTGGATTCAT | 4346 | | | | | | |
| Db | 2452 | CCATGGCGATCGCTGTTGCCGAATATCATGGTGGAAATGGCCGCTTTTCTGGATTCAT | 2511 | | | | | | |
| Qy | 4347 | CGACTGTGGCGGCTGGTGTGGCGGACCGCTATCAGGACATAGCTGTGGCTACCCGTGA | 4406 | | | | | | |
| Db | 2512 | CGACTGTGGCGGCTGGTGTGGCGGACCGCTATCAGGACATAGCTGTGGCTACCCGTGA | 2571 | | | | | | |
| Qy | 4407 | TATTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCTCGTGTTCATCGGTATCGC | 4466 | | | | | | |
| Db | 2572 | TATTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCTCGTGTTCATCGGTATCGC | 2631 | | | | | | |
| Qy | 4467 | CGCTCCCGATTCGACGCGCATCGCTTCTATCGCTTCTTGAACAGGTTCTTCTGAGCGGG | 4526 | | | | | | |
| Db | 2632 | CGCTCCCGATTCGACGCGCATCGCTTCTATCGCTTCTTGAACAGGTTCTTCTGAGCGGG | 2691 | | | | | | |
| Qy | 4527 | ACTCTGGGGT-----TCGATAAAATAAAGATTTTATTAGTCT | 4565 | | | | | | |
| Db | 2692 | ACTCTGGGGTTCGTGAGAAAGCTTGGGCCCATCGATAAATAAAGATTTTATTAGTCT | 2751 | | | | | | |
| Qy | 4566 | CCAGAAAGGGGGAAATGAAGAAGCCACCTGTAGGTTTGGCAGCTAGCTTAAGTAAC | 4625 | | | | | | |
| Db | 2752 | CCAGAAAGGGGGAAATGAAGAAGCCACCTGTAGGTTTGGCAGCTAGCTTAAGTAAC | 2811 | | | | | | |

| | | | |
|----|------|--|------|
| Qy | 4626 | GCCATTTTGAAGGCATGGAAAATACATAACTGAGAATAGAGAAGTTTCAGATCAAGGTC | 4685 |
| Db | 2812 | GCCATTTTGAAGGCATGGAAAATACATAACTGAGAATAGAGAAGTTTCAGATCAAGGTC | 2871 |
| Qy | 4686 | AGGAACAGATGGAAACAGCTGAATATGGGCCCAACAGAGATATCTGTGTAAAGCAGTTCTGT | 4745 |
| Db | 2872 | AGGAACAGATGGAAACAGCTGAATATGGGCCCAACAGAGATATCTGTGTAAAGCAGTTCTGT | 2931 |
| Qy | 4746 | CCCCGGCTCAGGGCCAGAAACAGATGGAACAGCTGAATATGGGCCCAACAGAGATATCTGT | 4805 |
| Db | 2932 | CCCCGGCTCAGGGCCAGAAACAGATGGAACAGCTGAATATGGGCCCAACAGAGATATCTGT | 2991 |
| Qy | 4806 | GTTAAGCAGTTCTTCCGCCCTCAGGGCCAAAGAACAGATGGAATGGTCCCGAGATGCGGTCCAG | 4865 |
| Db | 2992 | GTTAAGCAGTTCTTCCGCCCTCAGGGCCAAAGAACAGATGGAATGGTCCCGAGATGCGGTCCAG | 3051 |
| Qy | 4866 | CCCTCAGCAGTTTCTAGAGAACCATCAGATGTTTTCAGGGTGCCCCCAAGGACCTTGAATG | 4925 |
| Db | 3052 | CCCTCAGCAGTTTCTAGAGAACCATCAGATGTTTTCAGGGTGCCCCCAAGGACCTTGAATG | 3111 |
| Qy | 4926 | ACCTGTGCTTATTTGAACCTAACCATCAGTTTCGCTTCTCGCTTCTGTTCGCGCGCTTC | 4985 |
| Db | 3112 | ACCTGTGCTTATTTGAACCTAACCATCAGTTTCGCTTCTCGCTTCTGTTCGCGCGCTTC | 3171 |
| Qy | 4986 | TGCTCCCGCAGCTCAATAAAGAGAGCCCAACACCCCTCACCTCGGGGCGCCAGTCCCTCGAT | 5045 |
| Db | 3172 | TGCTCCCGCAGCTCAATAAAGAGAGCCCAACACCCCTCACCTCGGGGCGCCAGTCCCTCGAT | 3231 |
| Qy | 5046 | TGACTGAGTCGCGGGGTACCCGCTGATCAATAAACCCTCTTTCAGTTTGCATCCGACTT | 5105 |
| Db | 3232 | TGACTGAGTCGCGGGGTACCCGCTGATCAATAAACCCTCTTTCAGTTTGCATCCGACTT | 3291 |
| Qy | 5106 | GTGCTCTCGCTGTTCTTGGGAGGCTCTCTCTGAGTGATGACTACCCGTACGCGGGGG | 5165 |
| Db | 3292 | GTGCTCTCGCTGTTCTTGGGAGGCTCTCTCTGAGTGATGACTACCCGTACGCGGGGG | 3351 |
| Qy | 5166 | TCCTTCATTGG 5177 | |
| Db | 3352 | TCCTTCATTGG 3363 | |

RESULT 9
US-08-462-859A-6
; Sequence 6, Application US/08462859A
; Patent No. 5652092
; GENERAL INFORMATION:
; APPLICANT: Jacobsen, J. S.
; APPLICANT: Vitek, M. P.
; TITLE OF INVENTION: No. 5652092el Amyloid Precursor and Method of
; TITLE OF INVENTION: Using Same to Access Agents Which Down-Regulate Formation
; TITLE OF INVENTION: of B-Amyloid Peptide
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: American Cyanamid Company
; STREET: One Cyanamid Plaza
; CITY: Wayne
; STATE: New Jersey
; COUNTRY: United States
; ZIP: 07470-8426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/462,859A
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Barnhard, Elizabeth M.
; REGISTRATION NUMBER: 31,088
; REFERENCE/DOCKET NUMBER: 31,844-04

TELECOMMUNICATION INFORMATION:

TELEPHONE: (201)831-3246

TELEFAX: (201)831-3305

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 8591 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: circular

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 2393..3868

US-08-462-859A-6

Query Match 28.0%; Score 1448.4; DB 2; Length 8591;

Best Local Similarity 99.5%; Pred. No. 0;

Matches 1474; Conservative 0; Mismatches 6; Indels 2; Gaps 2;

QY 3062 TCCCTCAGGATATAGTAGTTTCGCTTTTGATAGGAGGGGAAATGATGCTTATGCA 3121
DB 4931 TCCCTCAGGATATAGTAGTTTCGCTTTTGATAGGAGGGGAAATGATGCTTATGCA 4990
QY 3122 ATACTCTTGATGCTTGCACATGCTACGATGAGTTAGCAACATGCTTACAGGAGAG 3181
DB 4991 ATACCTTGATGCTTGCACATGCTACGATGAGTTAGCAACATGCTTACAGGAGAG 5050
QY 3182 AAAAAGCACCGTGATGCTGCAATGCTGGAAGTAAGGTGCTGACGATCGTGCCTTATTAGGA 3241
DB 5051 AAAAAGCACCGTGATGCTGCAATGCTGGAAGTAAGGTGCTGACGATCGTGCCTTATTAGGA 5110
QY 3242 AGGCAACAGACGGGTCTGACATGATGGAAGCAACCACTGATGCTGCAATGCAAGAGAT- 3300
DB 5111 AGGCAACAGACGGGTCTGACATGATGGAAGCAACCACTGATGCTGCAATGCAAGAGAT 5170
QY 3301 ATTCTATTAGTGCCTAGCTCGATACAGCAACCGCCATTGACCATTCACCAATTGGT 3360
DB 5171 ATTCTATTAGTGCCTAGCTCGATACAGCAACCGCCATTGACCATTCACCAATTGGT 5230
QY 3361 GTGCACCTTCCAAAGCTTCAAGCTGCGCAAGCACTCAGGCGCAAGGGCTGCTAAAGGAAG 3420
DB 5231 GTGCACCTTCCAAAGCTTCAAGCTGCGCAAGCACTCAGGCGCAAGGGCTGCTAAAGGAAG 5290
QY 3421 CGGAACACGTAGAAGCCAGTCCGAGAAACCGTGTGACCCCGGATGAATGTGAGCTAC 3480
DB 5291 CGGAACACGTAGAAGCCAGTCCGAGAAACCGTGTGACCCCGGATGAATGTGAGCTAC 5350
QY 3481 TGGGCTATCTGGACAGGGAACCGCAAGCCAAAGAGAGAGAGTAGTTGAGTGGG 3540
DB 5351 TGGGCTATCTGGACAGGGAACCGCAAGCCAAAGAGAGAGAGTAGTTGAGTGGG 5410
QY 3541 CTTACATGGCGATAGCTAGCTGCGCGGCTTTTATGGACAGCAAGCAACCGGAATTGCCA 3600
DB 5411 CTTACATGGCGATAGCTAGCTGCGCGGCTTTTATGGACAGCAAGCAACCGGAATTGCCA 5470
QY 3601 GCTGGGCGCGCTCTGGTAAGTTGGGAAGCCCTGCAAGTAAGTAAGTGGCTTTCTTG 3660
DB 5471 GCTGGGCGCGCTCTGGTAAGTTGGGAAGCCCTGCAAGTAAGTAAGTGGCTTTCTTG 5530
QY 3661 CCGCAAGGATCTATGCGCGAGGGATCAAGATCTGATCAAGACAGGATAGGATCG 3720
DB 5531 CCGCAAGGATCTATGCGCGAGGGATCAAGATCTGATCAAGACAGGATAGGATCG 5590
QY 3721 TTTCCGATGATTGAACAGATGAGTTGACCCAGGTTCTCCGGCGCTTGGGTGGAGAGG 3780
DB 5591 TTTCCGATGATTGAACAGATGAGTTGACCCAGGTTCTCCGGCGCTTGGGTGGAGAGG 5650
QY 3781 CTATTGGCTATGACTGGGCAACACAGACATCGGCTGCTGATGCGCGCTGTTCCGG 3840
DB 5651 CTATTGGCTATGACTGGGCAACACAGACATCGGCTGCTGATGCGCGCTGTTCCGG 5710
QY 3841 CTGTACGCGAGGGCGCGCGGTTCTTTTGTCAAGACCGACCTGTGCGGTGCGCTGAAT 3900

DB 5711 CTGTACGCGAGGGCGCGCGGTTCTTTTGTCAAGACCGACCTGTGCGGTGCGCTGAAT 5770
QY 3901 GAACTGACGACGAGGCGCGGCTATCGTGGCTGGCCACGACGGGGGTTCTTTGCGCA 3960
DB 5771 GAACTGACGACGAGGCGCGGCTATCGTGGCTGGCCACGACGGGGGTTCTTTGCGCA 5830
QY 3961 GCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGACCTGGCTGCTATTGGGCGAAGTGGCG 4020
DB 5831 GCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGACCTGGCTGCTATTGGGCGAAGTGGCG 5890
QY 4021 GGGCAGGATCTCCCTGTCATCTCACCTTCTCTCCGAGAAAGTATCCATCATGCTGAT 4080
DB 5891 GGGCAGGATCTCCCTGTCATCTCACCTTCTCTCCGAGAAAGTATCCATCATGCTGAT 5950
QY 4081 GCATGCGGGGCTGCATAGCTTGTATCGGCTACCTGCGCATTCGACCAACCAAGCGAAA 4140
DB 5951 GCATGCGGGGCTGCATAGCTTGTATCGGCTACCTGCGCATTCGACCAACCAAGCGAAA 6010
QY 4141 CATGCGATCGAGCGGACGATCTCGGATGGAAGCCGCTTGTGCGATCAGGATGATCTG 4200
DB 6011 CATGCGATCG-GCGAGCACGATCTCGGATGGAAGCCGCTTGTGCGATCAGGATGATCTG 6069
QY 4201 GACGAAGACATCAGGGGCTCGCGCCAGCGCAACTGTTGCGCAGGCTCAAGCGCGCATG 4260
DB 6070 GACGAAGACATCAGGGGCTCGCGCCAGCGCAACTGTTGCGCAGGCTCAAGCGCGCATG 6129
QY 4261 CCCGACGCGGAGGATCTGCTGACCCCATGCGCGATGCTGCTTGCAGCAATATCATGCTG 4320
DB 6130 CCCGACGCGGAGGATCTGCTGACCCCATGCGCGATGCTGCTTGCAGCAATATCATGCTG 6189
QY 4321 GAAATGCGCGCTTTTCTGGAATTCATGCTGCGCGCTGGGTGGTGGCGGACCGCTAT 4380
DB 6190 GAAATGCGCGCTTTTCTGGAATTCATGCTGCGCGCTGGGTGGTGGCGGACCGCTAT 6249
QY 4381 CAGGACATAGCTTGGCTACCCGCTGATATGCTCAAGAGCTTGGCGGCAATGGGCTGAC 4440
DB 6250 CAGGACATAGCTTGGCTACCCGCTGATATGCTCAAGAGCTTGGCGGCAATGGGCTGAC 6309
QY 4441 CGCTTCTCGTCTTTACGGTATCGCGCTATCGCGCTCCCGATTCGACGCGCATGCTTATGCG 4500
DB 6310 CGCTTCTCGTCTTTACGGTATCGCGCTCCCGATTCGCGATTCGCGCATGCTTATGCG 6369
QY 4501 CTTCTTCAAGGATTTCTGAGCGGAGCTCTGGGGTTGATA 4542
DB 6370 CTTCTTCAAGGATTTCTGAGCGGAGCTCTGGGGTTGATA 6411

RESULT 10

US-08-462-859A-8

Sequence 8, Application US/08462859A

Patent No. 5652092

GENERAL INFORMATION:

APPLICANT: Jacobsen, J. S.

APPLICANT: Vittek, M. P.

TITLE OF INVENTION: No. 5652092el Amyloid Precursor and Method of

TITLE OF INVENTION: Using Same to Access Agents Which Down-Regulate Formation

TITLE OF INVENTION: of B-Amyloid Peptide

NUMBER OF SEQUENCES: 19

CORRESPONDENCE ADDRESS:

ADDRESSEE: American Cyanamid Company

STREET: One Cyanamid Plaza

CITY: Wayne

STATE: New Jersey

COUNTRY: United States

ZIP: 07470-8426

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/462,859A

FILING DATE: 05-JUN-1995

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICANT: US/08/123.659A
FILING DATE: 20-SEP-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Rosenblum, Anne M.
REGISTRATION NUMBER: 30,419
REFERENCE/DOCKET NUMBER: 31,844-01
TELECOMMUNICATION INFORMATION:
TELEPHONE: (518)475-0611
TELEFAX: (518)475-0619
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 8591 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 2393..3868
US-08-123-659A-6

Query Match 28.0%; Score 1448.4; DB 2; Length 8591;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1474; Conservative 0; Mismatches 6; Indels 2; Gaps 2;

| | | | |
|----|------|--|------|
| Qy | 3062 | TCCCTCAGGATATAGTATGTTTCGTTTTCGATAGGAGGGGAAATAGTCTTATGCA | 3121 |
| Db | 4931 | TCCCTCAGGATATAGTATGTTTCGTTTTCGATAGGAGGGGAAATAGTCTTATGCA | 4990 |
| Qy | 3122 | ATATCTTTGTAGTCTTGCACATGTAACGATAGTATGACCAACATGCTTACAAAGGAG | 3181 |
| Db | 4991 | ATACATTTGTAGTCTTGCACATGTAACGATAGTATGACCAACATGCTTACAAAGGAG | 5050 |
| Qy | 3182 | AAAAAGCAACCGTGATGCGATGTTGGTGAAGTAAGTGTAGTCTGCTTATTAGGA | 3241 |
| Db | 5051 | AAAAAGCAACCGTGATGCGATGTTGGTGAAGTAAGTGTAGTCTGCTTATTAGGA | 5110 |
| Qy | 3242 | AGGCAACAGACGGGTCTGACATGATTTGGACGAACCACTGATTCGCGATTCGAGAT- | 3300 |
| Db | 5111 | AGGCAACAGACGGGTCTGACATGATTTGGACGAACCACTGATTCGCGATTCGAGATA | 5170 |
| Qy | 3301 | ATTGTATTAAAGTGTAGTCTGATACAGCAAAACCGCATTTGACCATTCACCAATTGGT | 3360 |
| Db | 5171 | ATTGTATTAAAGTGTAGTCTGATACCAATTAACCGCATTTGACCATTCACCAATTGGT | 5230 |
| Qy | 3361 | GTGCAACCTCCAAAGCTTCAAGTGTGCGCAAGCACTCAGGGCGCAAGGGCTGCTAAAGGAAG | 3420 |
| Db | 5231 | GTGCAACCTCCAAAGCTTCAAGTGTGCGCAAGCACTCAGGGCGCAAGGGCTGCTAAAGGAAG | 5290 |
| Qy | 3421 | CGGAACAGTAAAGCCAGTCCGAGAAAACGGTGTGACCCCGATGAATGTGAGTAC | 3480 |
| Db | 5291 | CGGAACAGTAAAGCCAGTCCGAGAAAACGGTGTGACCCCGATGAATGTGAGTAC | 5350 |
| Qy | 3481 | TGGGCTATCTGACAAAGGAAACCGCAAGCGCAAGAGAAAGCAGTGTGAGTGGG | 3540 |
| Db | 5351 | TGGGCTATCTGACAAAGGAAACCGCAAGCGCAAGAGAAAGCAGTGTGAGTGGG | 5410 |
| Qy | 3541 | CTTACATGGCATAGTACGTGGCGGTTTATGACAGCAAGCAAGCAACCGGAATTGCA | 3600 |
| Db | 5411 | CTTACATGGCATAGTACGTGGCGGTTTATGACAGCAAGCAAGCAACCGGAATTGCA | 5470 |
| Qy | 3601 | GCTGGGCGCCCTCTGTGAAGTGTGGGAAGCCCTCAAAAGTAACTGGATGGCTTTCTTG | 3660 |
| Db | 5471 | GCTGGGCGCCCTCTGTGAAGTGTGGGAAGCCCTCAAAAGTAACTGGATGGCTTTCTTG | 5530 |
| Qy | 3661 | CGCCCAAGATCTGATGGCGGAGGAGTCAAGATCTGATCAAGACAGGATGAGGATCG | 3720 |
| Db | 5531 | CGCCCAAGATCTGATGGCGGAGGAGTCAAGATCTGATCAAGACAGGATGAGGATCG | 5590 |

| | | | |
|----|------|---|------|
| Qy | 3721 | TTTCGCATGATTGAACAAGATGGATTGCAACGACAGTTCTCCGGCGCTTGGGTGAGAGG | 3780 |
| Db | 5591 | TTTCGCATGATTGAACAAGATGGATTGCAACGACAGTTCTCCGGCGCTTGGGTGAGAGG | 5650 |
| Qy | 3781 | CTATTTCGGCTATGATCGGCGACAAACAGACAATCGGCTGCTCTGATGCGCGCTGTTCGG | 3840 |
| Db | 5651 | CTATTTCGGCTATGATCGGCGACAAACAGACAATCGGCTGCTCTGATGCGCGCTGTTCGG | 5710 |
| Qy | 3841 | CTGTACGCGACGGGCGCGGCTTCTTTTGTCAAGAACCAGACCTGTGCGGCTGCTGAT | 3900 |
| Db | 5711 | CTGTACGCGACGGGCGCGGCTTCTTTTGTCAAGAACCAGACCTGTGCGGCTGCTGAT | 5770 |
| Qy | 3901 | GAACCTGCAGGACGAGGCGCGCTATCGTGGCTGCGCACGACGAGGGGCTTCTTGGCA | 3960 |
| Db | 5771 | GAACCTGCAGGACGAGGCGCGCTATCGTGGCTGCGCACGACGAGGGGCTTCTTGGCA | 5830 |
| Qy | 3961 | GCTGTGCTCGACGTTGTCACTGAAGCGGAGGAGCTGGCTGCTATTGGGCGAAGTGGCG | 4020 |
| Db | 5831 | GCTGTGCTCGACGTTGTCACTGAAGCGGAGGAGCTGGCTGCTATTGGGCGAAGTGGCG | 5890 |
| Qy | 4021 | GGCAGGATCTCTGTCTCATCTCACCTTGTCTCTCCGAGAAAGTATCATCATGCTGAT | 4080 |
| Db | 5891 | GGCAGGATCTCTGTCTCATCTCACCTTGTCTCTCCGAGAAAGTATCATCATGCTGAT | 5950 |
| Qy | 4081 | GCATGCGGCGGCTGCATACGCTTGATCCGCTACCTGCCCATTCGACCAACAGCGAAA | 4140 |
| Db | 5951 | GCATGCGGCGGCTGCATACGCTTGATCCGCTACCTGCCCATTCGACCAACAGCGAAA | 6010 |
| Qy | 4141 | CATCGCATCGAGCAGACGATCTCGGATGGAAGCCGCTTTCGATCGACGATGATCTG | 4200 |
| Db | 6011 | CATCGCATCG - GCAGACGATCTCGGATGGAAGCCGCTTTCGATCGACGATGATCTG | 6069 |
| Qy | 4201 | GACGAAGACATCAGGGGCTCGCGCCAGCCGAACTGTTTCGCGAGGCTCAAGCGCGCATG | 4260 |
| Db | 6070 | GACGAAGACATCAGGGGCTCGCGCCAGCCGAACTGTTTCGCGAGGCTCAAGCGCGCATG | 6129 |
| Qy | 4261 | CCGACGCGGAGGATCTGCTGTCGACCCATGCGGATGCTGCTGCGGCAATATCATGCTG | 4320 |
| Db | 6130 | CCGACGCGGAGGATCTGCTGTCGACCCATGCGGATGCTGCTGCGGCAATATCATGCTG | 6189 |
| Qy | 4321 | GAAATGCGCGCTTTCTGGAATTCATCGACTGTGCGCGGCTGGTGTGCGGACCGCTAT | 4380 |
| Db | 6190 | GAAATGCGCGCTTTCTGGAATTCATCGACTGTGCGCGGCTGGTGTGCGGACCGCTAT | 6249 |
| Qy | 4381 | CAGGACATAGCGTTGGCTTACCCGCTGATATTCTGGAAGAGCTTGGCGGCAATGGGCTGAC | 4440 |
| Db | 6250 | CAGGACATAGCGTTGGCTTACCCGCTGATATTCTGGAAGAGCTTGGCGGCAATGGGCTGAC | 6309 |
| Qy | 4441 | CGCTTCTCGGCTTTTACGCTATCGCGCTCCCGATTCGACGCGCATCGCTTCTATCGC | 4500 |
| Db | 6310 | CGCTTCTCGGCTTTTACGCTATCGCGCTCCCGATTCGACGCGCATCGCTTCTATCGC | 6369 |
| Qy | 4501 | CTTCTTCAGGATTTCTTCTGAGCGGAGCTCTGGGGTTCGATA 4542 | |
| Db | 6370 | CTTCTTCAGGATTTCTTCTGAGCGGAGCTCTGGGGTTCGATA 6411 | |

RESULT 12

US-08-123-659A-8
; Sequence 8, Application US/08123659A
; Patent No. 5656477

GENERAL INFORMATION:

APPLICANT: Jacobsen, J. S.
APPLICANT: Vitok, M. P.
TITLE OF INVENTION: No. 5656477el Amyloid Precursor and Method of
; TITLE OF INVENTION: Using Same to Access Agents Which Down-Regulate Formation
; TITLE OF INVENTION: of B-Amyloid Peptide
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
ADDRESSEE: Anne Rosenblum
STREET: 163 Delaware Avenue, Suite 212
CITY: Delmar
STATE: New York

CORRESPONDENCE ADDRESS:
ADDRESS: American Cyanamid Company
STREET: One Campus Drive
CITY: Parsippany
STATE: New Jersey
COUNTRY: United States
ZIP: 07054
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/464,247A
FILING DATE: 05-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Barthard, Elizabeth M.
REGISTRATION NUMBER: 31,088
REFERENCE/DOCKET NUMBER: 31,844-03
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-683-2158
TELEFAX: 201-683-4117
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 8591 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 2393..3868
US-08-464-247A-6

| Query Match | 28.0%; | Score 1448.4; | DB 2; | Length 8591; |
|-----------------------|-----------------|---|-----------|--------------|
| Best Local Similarity | 99.5%; | Pred. No. 0; | | |
| Matches 1474; | Conservative 0; | Mismatches 6; | Indels 2; | Gaps 2; |
| Qy | 3062 | TCCCCTCAGGATATAGTAGTTTCGCTTTTCATATAGGAGGGGGAATGTAGTCTTATGCA | 3121 | |
| Db | 4931 | TCCCCTCAGGATATAGTAGTTTCGCTTTTGATAGGAGGGGGAATGTAGTCTTATGCA | 4990 | |
| Qy | 3122 | ATATCTTTGTAGTCTTCGAACATCGTAAACATGATTGAGCAACATGCTTTACAAGGAGAG | 3181 | |
| Db | 4991 | ATA CACTTTGTAGTCTTCGAACATCGTAAACATGATTGAGCAACATGCTTTACAAGGAGAG | 5050 | |
| Qy | 3182 | AAAAAGCACCGTGATCCGATTTGGTGGAAAGTAAGGTGGTACGATCGTGCTTTATTAGGA | 3241 | |
| Db | 5051 | AAAAAGCACCGTGATCCGATTTGGTGGAAAGTAAGGTGGTACGATCGTGCTTTATTAGGA | 5110 | |
| Qy | 3242 | AGGCAACAGACGGGTCTGACATGGATTGGACGAACCACTGAATTCGCATTTGCAGAGAT - | 3300 | |
| Db | 5111 | AGGCAACAGACAGGTCTGTGACATGGATTGGACGAACCACTGAATTCGCATTTGCAGAGATA | 5170 | |
| Qy | 3301 | ATTGTATTAAATGTCCTAGCTCGATACAGCAAAAGCCCAATTTGACCATTCACCACTTGGT | 3360 | |
| Db | 5171 | ATTGTATTTAAATGTCCTAGCTCGATACAAATAAACGCCAATTTGACCATTCACCACTTGGT | 5230 | |
| Qy | 3361 | GTGCACCTCCAAAGCTTCACGCTGCCGCAAGCACTCAGGGCGCAAGGCTCTTAAAGGAAG | 3420 | |
| Db | 5231 | GTGCACCTCTCAGCTTCACGCTGCCGCAAGCACTCAGGGCGCAAGGCTCTTAAAGGAAG | 5290 | |
| Qy | 3421 | CGGAACACGTAGAAAGCCAGTCCCGCAGAAACGGTCTGACCCCGGATGAATGTCAAGTAC | 3480 | |
| Db | 5291 | CGGAACACGTAGAAAGCCAGTCCCGCAGAAACGGTCTGACCCCGGATGAATGTCAAGTAC | 5350 | |
| Qy | 3481 | TGGGCTATCTGGAACAAGGGAAAAACGCAAGCGCAAGAGAAAGCAGGTAGCTTGCAGTGGG | 3540 | |
| Db | 5351 | TGGGCTATCTGGAACAAGGGAAAAACGCAAGCGCAAGAGAAAGCAGGTAGCTTGCAGTGGG | 5410 | |
| Qy | 3541 | CTTACATGGCGATAGCTAGACTGGGCGGTTTTATTGGA CAGCAACGCAACCGGAATTGSCCA | 3600 | |

RESULT 14
US-08-464-247A-8
; Sequence 8, Application US/08464247A
; Patent No. 5693478
; GENERAL INFORMATION:
; APPLICANT: Jacobsen, J. S.

APPLICANT: Vitek, M. P.
TITLE OF INVENTION: No. 5693478el Amyloid Precursor and Method of
TITLE OF INVENTION: Using Same to Access Agents Which Down-Regulate Formation
TITLE OF INVENTION: of B-Amyloid Peptide
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESS: American Cyanamid Company
STREET: One Campus Drive
CITY: Parsippany
STATE: New Jersey
COUNTRY: United States
ZIP: 07054
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/464,247A
FILING DATE: 05-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Barnhard, Elizabeth M.
REGISTRATION NUMBER: 31,088
REFERENCE/DOCKET NUMBER: 31,844-03
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-683-2158
TELEFAX: 201-683-4117
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 8591 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 2393..3853
US-08-464-247A-8

Query Match 28.0%; Score 1448.4; DB 2; Length 8591;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1474; Conservative 0; Mismatches 6; Indels 2; Gaps 2;
QY 3062 TCCCTCAGGATATAGTATTCGCTTTTTCATAGGAGGGGAAATAGTCTTATGCA 3121
DB 4931 TCCCTCAGGATATAGTATTCGCTTTTTCATAGGAGGGGAAATAGTCTTATGCA 4990
QY 3122 ATACTCTTGTAGTCTTGCAACATGTAACGATAGTATAGCAACATGCTTACAAGGAG 3181
DB 4991 ATACACTTGTAGTCTTGCAACATGTAACGATAGTATAGCAACATGCTTACAAGGAG 5050
QY 3182 AAAAAGCACCGTGCAATGCCGATTTGGTGAAGTAAGGTGGTACGATCGTCCCTTATTAGGA 3241
DB 5051 AAAAAGCACCGTGCAATGCCGATTTGGTGAAGTAAGGTGGTACGATCGTCCCTTATTAGGA 5110
QY 3242 AGGCAACAGACCGGTCTGACATGGAATTCGACGAACCACTGAAATCCGCAATTCGACAGAT- 3300
DB 5111 AGGCAACAGACCGGTCTGACATGGAATTCGACGAACCACTGAAATTCGCAATTCGACAGATA 5170
QY 3301 ATTGTATTATAGTCTAGCTAGCTACAGCAAGCAAGCCATTTGACCATTCACCACTGGT 3360
DB 5171 ATTGTATTATAGTCTAGCTAGCTACAGCAAGCAAGCCATTTGACCATTCACCACTGGT 5230
QY 3361 GTGCACCTCCCAAGCTTACGCTGCCGCAAGCACTCAGGCGCCCAAGGGCTGCTAAAGGAAG 3420
DB 5231 GTGCACCTCCCTAGCTTACGCTGCCGCAAGCACTCAGGCGCCCAAGGGCTGCTAAAGGAAG 5290
QY 3421 CGGAACACGTAGAAAGCCAGTCCGCAAGAAAGCGGTCTGACCCCGGATGAATGTGAGCTAC 3480
DB 5291 CGGAACACGTAGAAAGCCAGTCCGCAAGAAAGCGGTCTGACCCCGGATGAATGTGAGCTAC 5350
QY 3481 TGGGCTATCTGACAAAGGGAACCGCAAGCGCAAGAGCAAGGAGTAGCTTGCAGTGGG 3540

DB 5351 TGGGCTATCTGGAACAAGGGAACCGCAAGCGCAAGAGAAAGCAGGTAGCTTGCAGTGGG 5410
QY 3541 CTTACATGGCGATAGCTAGACTGGCGGTTTATGGACAGCAAGCGAACCGGAATTCGCA 3600
DB 5411 CTTACATGGCGATAGCTAGACTGGCGGTTTATGGACAGCAAGCGAACCGGAATTCGCA 5470
QY 3601 GCTGGGGCGCCCTCTGTTAGGTTGGGAAGCCCTGCAAAAGTAAACTGATGGCTTTCTTG 3660
DB 5471 GCTGGGGCGCCCTCTGTTAGGTTGGGAAGCCCTGCAAAAGTAAACTGATGGCTTTCTTG 5530
QY 3661 CCGCAAGGATCTGATCGCGCAGGGGATCAAGATCTGATCAAGAGACAGGATGAGATCG 3720
DB 5531 CCGCAAGGATCTGATCGCGCAGGGGATCAAGATCTGATCAAGAGACAGGATGAGATCG 5590
QY 3721 TTTTCGATGATTGAAAGATGGAATTCGACGCGAGTTCTCCGCGGCTTGGGTGGAGAGG 3780
DB 5591 TTTTCGATGATTGAAAGATGGAATTCGACGCGAGTTCTCCGCGGCTTGGGTGGAGAGG 5650
QY 3781 CTATTTCGGCTATGACTGGGCAACAACAGCAATCGGCTGCTGATGCCCGCTGTTCCGG 3840
DB 5651 CTATTTCGGCTATGACTGGGCAACAACAGCAATCGGCTGCTGATGCCCGCTGTTCCGG 5710
QY 3841 CTGTACGCGCAGGGCGCCCGGTTCTTTTGTCAAGACCGACCTGTCCGGTGCCTGAAT 3900
DB 5711 CTGTACGCGCAGGGCGCCCGGTTCTTTTGTCAAGACCGACCTGTCCGGTGCCTGAAT 5770
QY 3901 GAATGCGAGACGAGGCGCGGCTATGTTGGTGGCCACGACGCGGGGTTCTTTCGCA 3960
DB 5771 GAATGCGAGACGAGGCGCGGCTATGTTGGTGGCCACGACGCGGGGTTCTTTCGCA 5830
QY 3961 GCTGTGCTCGACGTTGCTACCTGAAGCGGGAAGGACTGGCTGCTATTGGGCGAAGTGC 4020
DB 5831 GCTGTGCTCGACGTTGCTACCTGAAGCGGGAAGGACTGGCTGCTATTGGGCGAAGTGC 5890
QY 4021 GGGCAGGATCTCTGTGATCTCCTCTGCTCCGAGAAAGTATCCATCATGGCTGAT 4080
DB 5891 GGGCAGGATCTCTGTGATCTCCTCTGCTCCGAGAAAGTATCCATCATGGCTGAT 5950
QY 4081 GCAATGGGGCGGCTGCAATGCTTGTGATCCGGCTACCTGCGCATTCGACCAACAGCGAA 4140
DB 5951 GCAATGGGGCGGCTGCAATGCTTGTGATCCGGCTACCTGCGCATTCGACCAACAGCGAA 6010
QY 4141 CATCGCATCGAGGCGACGACTCTCGATGGAAGCGGCTTTGTCGATCAGGATGATCTG 4200
DB 6011 CATCGCATCG - GCAGACGATCTCGATGGAAGCGGCTTTGTCGATCAGGATGATCTG 6069
QY 4201 GACGAAGAGCATCAGGGGCTCGCGCCAGCCGAACTGTTCCGACGGCTCAAGGCGCGCATG 4260
DB 6070 GACGAAGAGCATCAGGGGCTCGCGCCAGCCGAACTGTTCCGACGGCTCAAGGCGCGCATG 6129
QY 4261 CCGAGCGGAGGATCTGCTGTGATGCCCATGCGGATGCTGCTGCGCAATATCATGGTG 4320
DB 6130 CCGAGCGGAGGATCTGCTGTGATGCCCATGCGGATGCTGCTGCGCAATATCATGGTG 6189
QY 4321 GAAATCGCGCTTTCTGGAATTCATCGACTGTGGCGGCTGGGTGGGCGGACCGCTAT 4380
DB 6190 GAAATCGCGCTTTCTGGAATTCATCGACTGTGGCGGCTGGGTGGGCGGACCGCTAT 6249
QY 4381 CAGGACATAGCGTTGGCTAGCTGATATTTGCTGAAGAGCTTGGCGGCGAATGGGCTGAC 4440
DB 6250 CAGGACATAGCGTTGGCTAGCTGATATTTGCTGAAGAGCTTGGCGGCGAATGGGCTGAC 6309
QY 4441 CGCTTCCTGCTGCTTTACGGTATGCGCGCTCCCGAATTCGAGCGCATCGCTTCTATCGC 4500
DB 6310 CGCTTCCTGCTGCTTTACGGTATGCGCGCTCCCGAATTCGAGCGCATCGCTTCTATCGC 6369
QY 4501 CTTCTTCAGGATCTTTCTGAGCGGAGCTCTGGGGTTTCGATA 4542
DB 6370 CTTCTTCAGGATCTTTCTGAGCGGAGCTCTGGGGTTTCGATA 6411

RESULT 15

US-08-464-248A-6
 ; Sequence 6, Application US/08464248A
 ; Patent No. 5703209
 ; GENERAL INFORMATION:
 ; APPLICANT: Jacobsen, J. S.
 ; APPLICANT: Vittek, M. P.
 ; TITLE OF INVENTION: No. 5703209el Anyloid Precursor and Method of
 ; TITLE OF INVENTION: Using Same to Access Agents Which Down-Regulate Formation
 ; TITLE OF INVENTION: of B-Amyloid Peptide
 ; NUMBER OF SEQUENCES: 19
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: American Cyanamid Company
 ; STREET: One Cyanamid Plaza
 ; CITY: Wayne
 ; STATE: New Jersey
 ; COUNTRY: United States
 ; ZIP: 07470-8426
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/464,248A
 ; FILING DATE: 05-JUN-1995
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Barnhard, Elizabeth M.
 ; REGISTRATION NUMBER: 31,088
 ; REFERENCE/DOCKET NUMBER: 31,844-02
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (201)831-3246
 ; TELEFAX: (201)831-3305
 ; INFORMATION FOR SEQ ID NO: 6:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 8591 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: circular
 ; MOLECULE TYPE: cDNA
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 2393..3868
 ; US-08-464-248A-6

Query Match 28.0%; Score 1448.4; DB 2; Length 8591;
 Best Local Similarity 99.5%; Pred. No. 0;
 Matches 1474; Conservative 0; Mismatches 6; Indels 2; Gaps 2;
 Qy 3062 TCCCTCAGGATATAGTAGTTTCGCTTTTGATAGGAGGGGAAATGTAGTCTTATGCA 3121
 Db 4931 TCCCTCAGGATATAGTAGTTTCGCTTTTGATAGGAGGGGAAATGTAGTCTTATGCA 4990
 Qy 3122 ATACTCTGTAGTCTTGCAACATGTAACGATAGTAGTACCAATGCTTCAAGAGAG 3181
 Db 4991 ATACACTGTAGTCTTGCAACATGTAACGATAGTAGTACCAATGCTTCAAGAGAG 5050
 Qy 3182 AAAAAGCACCTGCATGCCGATGCTGGAAGTAAGGTGTTACGATCGTGCCTTATTAGGA 3241
 Db 5051 AAAAAGCACCTGCATGCCGATGCTGGAAGTAAGGTGTTACGATCGTGCCTTATTAGGA 5110
 Qy 3242 AGGCAACAGACGGGTCTGACATGGAATTTGGAAGAACCACTGAAATTCGCAATTTGACAGAT - 3300
 Db 5111 AGGCAACAGACGGGTCTGACATGGAATTTGGAAGAACCACTGAAATTCGCAATTTGACAGATA 5170
 Qy 3301 ATTGATTTAAGTCTAGTCTGATACAGCAACAGCAATTTGACCAATTTGACCAATTTGGT 3360
 Db 5171 ATTGATTTAAGTCTAGTCTGATACCAATTTGACCAATTTGACCAATTTGACCAATTTGGT 5230
 Qy 3361 GTGCACCTCCAGCTTCAAGCTGCGCAAGCACTCAGGCGCAAGGGCTGCTTAAGAGAG 3420
 Db 5231 GTGCACCTCCAGCTTCAAGCTGCGCAAGCACTCAGGCGCAAGGGCTGCTTAAGAGAG 5290

Qy 3421 CGGAACACGTAGAAAGCCAGTCCGCAGAAACGGTGTGACCCCGGATGAATGTGAGCTAC 3480
 Db 5291 CGGAACACGTAGAAAGCCAGTCCGCAGAAACGGTGTGACCCCGGATGAATGTGAGCTAC 5350
 Qy 3481 TGGGCTATCTGGACAAGGGAAACGCAAGCGCAAGAGAAAGCAAGTGTGAGTGTGAGTGGG 3540
 Db 5351 TGGGCTATCTGGACAAGGGAAACGCAAGCGCAAGAGAAAGCAAGTGTGAGTGTGAGTGGG 5410
 Qy 3541 CTTCATATGGGATAGTAGTGGGCGGTTTATATGGACAGCAAGCAAGCAAGCAAGTGTGCA 3600
 Db 5411 CTTCATATGGGATAGTAGTGGGCGGTTTATATGGACAGCAAGCAAGCAAGTGTGCA 5470
 Qy 3601 GCTGGGCGGCTCTGTAGAGTTGGGAAGCCCTGCAAAATGAATGATGATGCTTCTTG 3660
 Db 5471 GCTGGGCGGCTCTGTAGAGTTGGGAAGCCCTGCAAAATGAATGATGATGCTTCTTG 5530
 Qy 3661 CCGCAAGGATCTGATGGCGCAGGGATCAAGATCTGATCAAGAGACAGGATGAGGATCG 3720
 Db 5531 CCGCAAGGATCTGATGGCGCAGGGATCAAGATCTGATCAAGAGACAGGATGAGGATCG 5590
 Qy 3721 TTTGCGCATGATTGAACAAGATGGATTGACGACAGTCTCGGGCGCTTGGGTGGAGAGG 3780
 Db 5591 TTTGCGCATGATTGAACAAGATGGATTGACGACAGTCTCGGGCGCTTGGGTGGAGAGG 5650
 Qy 3781 CTATTGCGCTATGATCGGCGACAAACAGCAATCGGCTGCTGATGCGCGCGTGTTCGG 3840
 Db 5651 CTATTGCGCTATGATCGGCGACAAACAGCAATCGGCTGCTGATGCGCGCGTGTTCGG 5710
 Qy 3841 CTGTGCGCAGGGCGGCGGCTTCTTTTGTCAAGACCGACCTGCTCGGTGCGCTTGAAT 3900
 Db 5711 CTGTGCGCAGGGCGGCGGCTTCTTTTGTCAAGACCGACCTGCTCGGTGCGCTTGAAT 5770
 Qy 3901 GAATGCGACGAGGCGGCGGCTATCGTGGCTGCGGCGGCGGCTTCTTGGGCA 3960
 Db 5771 GAATGCGACGAGGCGGCGGCTATCGTGGCTGCGGCGGCGGCTTCTTGGGCA 5830
 Qy 3961 GCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGACTGGGTGCTATTGGGCGAAGTGGC 4020
 Db 5831 GCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGACTGGGTGCTATTGGGCGAAGTGGC 5890
 Qy 4021 GGGCAGGATCTCTGCTCATCTCACCTTGTCTTCCGCGAGAGATTCATCATGCTGAT 4080
 Db 5891 GGGCAGGATCTCTGCTCATCTCACCTTGTCTTCCGCGAGAGATTCATCATGCTGAT 5950
 Qy 4081 GCAATGCGGCGGTGCTATGATCGGCTACCTGCGCATTCGACACCAAGCGGAA 4140
 Db 5951 GCAATGCGGCGGTGCTATGATCGGCTACCTGCGCATTCGACACCAAGCGGAA 6010
 Qy 4141 CATCGCATCG-GCGAGCACGTACTCGGATGGAAGCGGCTTGTGCGATCAGGATGATCTG 4200
 Db 6011 CATCGCATCG-GCGAGCACGTACTCGGATGGAAGCGGCTTGTGCGATCAGGATGATCTG 6069
 Qy 4201 GACGAGAGCATCAGGGGCTCGGCGCAGCCGAACTGTTCCGAGGCTCAAGGGCGGCA 4260
 Db 6070 GACGAGAGCATCAGGGGCTCGGCGCAGCCGAACTGTTCCGAGGCTCAAGGGCGGCA 6129
 Qy 4261 CCGGACGCGGAGGATCTGCTGACCATGCGGATGCGCTTGTGCGGAAATATCATGGT 4320
 Db 6130 CCGGACGCGGAGGATCTGCTGACCATGCGGATGCGCTTGTGCGGAAATATCATGGT 6189
 Qy 4321 GAAATGCGGCTTCTGGAATCATGATGCTGGCGGCTGGGTGTGGCGGACCGCTAT 4380
 Db 6190 GAAATGCGGCTTCTGGAATCATGATGCTGGCGGCTGGGTGTGGCGGACCGCTAT 6249
 Qy 4381 CAGGACATAGCGTTGGCTACCGCTGATTTGCTGAAGAGCTTGGCGGCGAATGGGCTGAC 4440
 Db 6250 CAGGACATAGCGTTGGCTACCGCTGATTTGCTGAAGAGCTTGGCGGCGAATGGGCTGAC 6309
 Qy 4441 CGCTTCTCTGCTTTACGCTATCGCGCTCCGATTCGAGCGCATTCGCTTCTTATCGC 4500
 Db 6310 CGCTTCTCTGCTTTACGCTATCGCGCTCCGATTCGAGCGCATTCGCTTCTTATCGC 6369
 Qy 4501 CTTCTTGACGAGTTCTTCTGAGGGGACTCTGGGGTTTCGATA 4542

Db 6370 CTTCTTGACGAGTTCTTCTGAGCGGACTCTGGGGTTCGAAA 6411
|||||

Search completed: February 27, 2006, 08:49:42
Job time : 893 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 27, 2006, 08:49:57 ; Search time 6646 Seconds
(without alignments)
3421.213 Million cell updates/sec

Title: US-10-618-570-2

Perfect score: 2059

Sequence: 1 MAQWDDPPDQEDTDSCTES.....DWKATRVGINIFRLRTQKE 400

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/abs/ABSSWEB.apool/US10618570/runat 27022006 063854 13050/app query.fasta_1
-DB=GenEmbl -OPMT=fastap -SUPFIX=rge -MINMATCH=0.1 -LOPFCI=0 -LOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTPMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs02h
-USER=US10618570 @CGN 1 1 4939 @runat 27022006 063854 13050 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:*

1: gb_ba:*
2: gb_in:*
3: gb_env:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pt:*
9: gb_pro:*
10: gb_sts:*
11: gb_sy:*
12: gb_un:*
13: gb_vi:*
14: gb_btg:*
15: gb_pli:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|----|--------------------|
| 1 | 2059 | 100.0 | 5177 | 6 | BD136828 Biotin-bi |
| 2 | 2059 | 100.0 | 5177 | 6 | AX018983 Sequence |
| 3 | 1381.5 | 67.1 | 1588 | 6 | AR365560 Sequence |

| 4 | 1381.5 | 67.1 | 1807 | 6 | AR365559 | Sequence |
|----|--------|------|--------|----|-----------|--------------------|
| 5 | 1381 | 67.1 | 1592 | 4 | BISCAVR1 | X51689 B.taurus mr |
| 6 | 1381 | 67.1 | 2135 | 4 | BISCAVR2 | X54183 B.taurus mr |
| 7 | 1042.5 | 50.6 | 1167 | 8 | AF037351 | AF037351 Homo sapi |
| 8 | 1042.5 | 50.6 | 1347 | 6 | E03268 | E03268 cDNA sequen |
| 9 | 1042.5 | 50.6 | 1347 | 6 | E05211 | E05211 DNA encodin |
| 10 | 1042.5 | 50.6 | 1347 | 6 | BD000149 | BD000149 Anthuman |
| 11 | 1042.5 | 50.6 | 1367 | 6 | CQ722728 | CQ722728 Sequence |
| 12 | 1042.5 | 50.6 | 1367 | 6 | I72357 | I72357 Sequence 3 |
| 13 | 1042.5 | 50.6 | 1367 | 6 | AR270697 | AR270697 Sequence |
| 14 | 1042.5 | 50.6 | 1367 | 8 | HUMPHSR2 | D90188 Homo sapien |
| 15 | 1042.5 | 50.6 | 2028 | 6 | E03267 | E03267 cDNA sequen |
| 16 | 1042.5 | 50.6 | 2028 | 6 | E05210 | E05210 DNA encodin |
| 17 | 1042.5 | 50.6 | 2028 | 6 | I72356 | I72356 Sequence 1 |
| 18 | 1042.5 | 50.6 | 2028 | 6 | BD000148 | BD000148 Anthuman |
| 19 | 1042.5 | 50.6 | 2028 | 8 | HUMPHSR1 | D90187 Homo sapien |
| 20 | 1042.5 | 50.6 | 2209 | 8 | BC063878 | BC063878 Homo sapi |
| 21 | 1042.5 | 50.6 | 3704 | 6 | AX281714 | AX281714 Sequence |
| 22 | 1039.5 | 50.5 | 2037 | 6 | I41352 | I41352 Sequence 1 |
| 23 | 994.5 | 48.3 | 1330 | 6 | AX928628 | AX928628 Sequence |
| 24 | 994.5 | 48.3 | 1330 | 6 | AX926677 | AX926677 Sequence |
| 25 | 994.5 | 48.3 | 1330 | 6 | RABMSRT1 | D13381 Oryctolagus |
| 26 | 987.5 | 48.0 | 1563 | 4 | RABMSRT1A | L11692 Oryctolagus |
| 27 | 987.5 | 48.0 | 1862 | 4 | RABMSRTIA | L11693 Oryctolagus |
| 28 | 819 | 39.8 | 1425 | 9 | AF203781 | AF203781 Mus muscu |
| 29 | 801 | 38.9 | 1065 | 9 | MUSMSRT2 | D13382 Mus sp. msn |
| 30 | 801 | 38.9 | 1430 | 9 | MUSMSRT1 | L04274 Mus musculu |
| 31 | 801 | 38.9 | 1572 | 9 | BC003814 | BC003814 Mus muscu |
| 32 | 801 | 38.9 | 1945 | 9 | MUSMSRT1I | L04275 Mus musculu |
| 33 | 710.5 | 34.5 | 257208 | 14 | AC157118 | AC157118 Bos tauru |
| 34 | 688 | 33.4 | 459 | 6 | C0830753 | C0830753 Sequence |
| 35 | 678.5 | 33.0 | 1746 | 6 | CS104130 | CS104130 Sequence |
| 36 | 678.5 | 33.0 | 1746 | 11 | AJ966780 | AJ966780 Synthetic |
| 37 | 678 | 32.9 | 604 | 5 | GGAVIR | X05343 Chicken mRN |
| 38 | 678 | 32.9 | 604 | 6 | A93651 | A93651 Sequence 3 |
| 39 | 678 | 32.9 | 604 | 6 | AR177253 | AR177253 Sequence |
| 40 | 678 | 32.9 | 604 | 6 | AR217937 | AR217937 Sequence |
| 41 | 677.5 | 32.9 | 873 | 6 | CS104134 | CS104134 Sequence |
| 42 | 586 | 28.5 | 897 | 6 | CS104131 | CS104131 Sequence |
| 43 | 586 | 28.5 | 897 | 11 | AJ616762 | AJ616762 Synthetic |
| 44 | 537 | 26.1 | 1119 | 10 | CHKAVID | L27818 Gallus gall |
| 45 | 523 | 25.4 | 1224 | 5 | GGA311647 | AJ311647 Gallus ga |

ALIGNMENTS

| RESULT 1 | BD136828 | BD136828 | 5177 bp | DNA | linear | PAT 18-SEP-2002 |
|------------|---|-----------------------------------|---------|-----|--------|-----------------|
| LOCUS | BD136828 | Biotin-binding receptor molecule. | | | | |
| DEFINITION | BD136828 | | | | | |
| ACCESSION | BD136828 | | | | | |
| VERSION | BD136828.1 | GI:23231773 | | | | |
| KEYWORDS | JP 2002504328-A/1. | | | | | |
| SOURCE | unidentified | | | | | |
| ORGANISM | unclassified. | | | | | |
| REFERENCE | 1 (bases 1 to 5177) | | | | | |
| AUTHORS | Herttua, S.V., Kulomaa, M., Lehtolainen, P., Marjomaki, V. and Airenne, K. | | | | | |
| TITLE | Biotin-binding receptor molecule | | | | | |
| JOURNAL | Patent: JP 2002504328-A 1 12-FEB-2002; | | | | | |
| COMMENT | ARC THERAPEUTICS LTD | | | | | |
| | OS Unidentified | | | | | |
| | PN JP 2002504328-A/1 | | | | | |
| | PD 12-FEB-2002 | | | | | |
| | PF 23-FEB-1999 JP 2000532517 | | | | | |
| | PP 23-FEB-1998 GB 980375.5, 24-JUN-1998 GB 9813653.4 PI | | | | | |
| | SPPO YLA HERTTUALA, MARKKU KULOMAA, PAULIINA LEHTOLAINEN, VARPPI MARJOMAKI, | | | | | |
| | PI KARI AJRENNNE | | | | | |
| | PC C12N15/09, A61K48/00, A61P43/00, C07K14/705//A61K38/00, C12N15/00, | | | | | |
| | PC A61K37/02 | | | | | |
| | CC Strandedness: Single; | | | | | |

Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 6 Gaps: 0

US-10-618-570-2 (1-400) x AX018983 (1-5177)

Qy 1 MetAlaGlnTTPAspAspPheProAspGlnGlnAspThrAspSerCysThrGluSer 20
 Db 1071 ATGGCAGAGTGGGATGACTTTCTCGATCAGCAAGAGGACACTGACAGCTGTACAGAGTCT 1130

Qy 21 VallysPheAspAlaAatgSerValThrAlaLeuLeuProProHisProLysAsnGlyPro 40
 Db 1131 GTGAAGTTCGATGCTCGCTCAGTCAGTCAGCTTTGCTTCTCCCATCTCTAAAAATGGGCCA 1190

Qy 41 ThrLeuGlnGluArgMetLysSerTyrLysThrAlaLeuLeuLeuLeuLeuVal 60
 Db 1191 ACTCTTCAAGAGAGGATGAAGTCTTATAAACTGCACTGATCACCTTTATCTCATTTGTG 1250

Qy 61 PheValValLeuValProIleIleGlyIleValAlaAlaGlnLeuLeuLeuLeuThr 80
 Db 1251 TTTGTAGTTCCTGCTGCCCATCATTTGGCATATGGCAGCTCAGCTCCTGAAATGGGAACG 1310

Qy 81 LysAsnCysThrValGlySerValAsnAlaAspIleSerProSerProGluGlyLysGly 100
 Db 1311 AAGAAATTCACAGGTTGGCTCAGTTAATGCAGATATATCTCCAGTCCGGAAGGCAAGGA 1370

Qy 101 AsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsnMet 120
 Db 1371 AATGCAGTGAAGATGAATGAGATTTGGAAGCTGTGATGGAACGCATGAGCAACATG 1430

Qy 121 GluSerArgIleGlnTyrLeuSerAspAsnGluAlaAsnLeuLeuLeuLeuAsnPhe 140
 Db 1431 GAAAGCAGAAATCCAGTATCTTTTCAGATTAATGAAGCCATCTCTCAGATGCTTAAGAAATTC 1490

Qy 141 GlnAsnPheSerIleThrThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsnSer 160
 Db 1491 CAAATTTTCAGATTAACACTGATCAAGATTAATGATGTTCTTTCCAGCTAAATCC 1550

Qy 161 LeuLeuSerSerIleGlnGluHisGluAsnIleIleGlyAspIleSerLysSerLeuVal 180
 Db 1551 TTACTTTCTCCATCCAGGAACATGAGAAATATATATAGGGATATCTCCAGCTATTAGTA 1610

Qy 181 GlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArgVal 200
 Db 1611 GGTCTGAACACCACTAGTACTTGAATTTGAGTATTTGAACACTCAATGGCAGAGTC 1670

Qy 201 GlnGluAsnAlaPheLysGlnGlnGluMetArgLysLeuGluArgIleTyrAsn 220
 Db 1671 CAAGAGATGCATTTAAACAACAAGAGAGATGCGTAATTTAGAGGAGCGGTATATACAAT 1730

Qy 221 AlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyrLeuGluGlnGluIleLys 240
 Db 1731 GCATCAGCAGAAATTAAGTCTCTAGATGAATAAACAAGTATATTTGGACAGGAATAAATA 1790

Qy 241 GlyGluMetLysLeuLeuAsnAsnIleThrAsnAspLeuArgLeuLysAspTyrGluHis 260
 Db 1791 GGGGAAATGAAACTGTTGAATAATATCACTAATGATCTGAGGCTGAAGGATGGGAACAT 1850

Qy 261 SerGlnThrLeuLysAsnIleThrLeuLeuGlnGlyAlaArgLysCysSerLeuThrGly 280
 Db 1851 TCTCAGACATTTGAAAATATACATTTACTCAAGGTGCCAAGATGCTCGCTGACTGGG 1910

Qy 281 LysTyrThrAsnAspLeuGlySerAsnMetThrIleGlyAlaValAsnSerArgGlyGlu 300
 Db 1911 AAATGGACCAACGATCTGGGCTCCACATGACCATCGGGCTGTGACAGCAGAGGTGAA 1970

Qy 301 PheThrGlyThrTyrIleThrAlaValThrAlaThrSerAsnGluIleLysGluSerPro 320
 Db 1971 TTCAAGCAGCCTACATCAGACCGGTAAACGCCACATCAATGAGATCAAGAGTCAACA 2030

Qy 321 LeuHisGlyThrGlnAsnThrIleAsnLysArgThrGlnProThrPheGlyPheThrVal 340
 Db 2031 CTGCATGGGACACAAACACCATCAACAAGAGGCCACCCAGCCACCTTTGGCTTCCCGTC 2090

Qy 341 AsnTyrLysPheSerGluSerThrThrValPheThrGlyGlnCysPheIleAspArgAsn 360
 Db 2091 AATTGGAAGTTTTCAGAGTCCACCAGTCTTTCACGGGCCAGTCTTTCATAGACAGGAAT 2150

Qy 361 GlyLysGluValLeuLeuLysThrMetTyrLeuLeuArgSerValAsnAspIleGlyAsp 380
 Db 2151 GGGGAAGGAGGTCCTGAAGACCATGTGCTGCTGGGTCAGATGTTAATGACATTTGGTAT 2210

Qy 381 AspTyrLysAlaThrArgValGlyIleAsnIlePheThrArgLeuArgThrGlnLysGlu 400
 Db 2211 GACTGGAAAGCTACAGGTCGGCATCAACATCTTCACTCGCTCGGCACACAGAGGAG 2270

RESULT 3
 AR365560 1588 bp DNA linear PAT 03-SEP-2003
 LOCUS Sequence 3 from patent US 5510466.
 DEFINITION
 AR365560
 ACCESSION
 AR365560.1 GI:34429331
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Unknown.
 Unclassified.
 REFERENCE
 1 (bases 1 to 1588)
 AUTHORS
 Krieger, M. and Kodama, T.
 TITLE
 Scavenger receptor protein and antibody thereto
 JOURNAL
 Patent: US 5510466-A 3 23-APR-1996;
 Massachusetts Institute of Technology; Cambridge, MA
 FEATURES
 Location/Qualifiers
 source
 1..1588
 /organism="unknown"
 /mol_type="genomic DNA"

ORIGIN

Alignment Scores:
 Pred. No.: 3,31e-89 Length: 1588
 Score: 1381.50 Matches: 307
 Percent Similarity: 79.4% Conservative: 24
 Best Local Similarity: 73.6% Mismatches: 66
 Query Match: 67.1% Indels: 22
 DB: 6 Gaps: 6

US-10-618-570-2 (1-400) x AR365560 (1-1588)

Qy 1 MetAlaGlnTTPAspAspPheProAspGlnGlnAspThrAspSerCysThrGluSer 20
 Db 4 ATGGCAGAGTGGGATGACTTTCTCGATCAGCAAGAGGACACTGACAGCTGTACAGAGTCT 63

Qy 21 VallysPheAspAlaAatgSerValThrAlaLeuLeuProProHisProLysAsnGlyPro 40
 Db 64 GTGAAGTTCGATGCTCGCTCAGTCAGACAGCTTTGCTTCTCCCATCTCTAAAAATGGGCCA 123

Qy 41 ThrLeuGlnGluArgMetLysSerTyrLysThrAlaLeuLeuLeuLeuLeuVal 60
 Db 124 ACTCTTCAAGAGAGGATGAAGTCTTATAAACTGCACTGATCACCTTTATCTCATTTGTG 183

Qy 61 PheValValLeuValProIleIleGlyIleValAlaAlaGlnLeuLeuLeuLysTyrGluThr 80
 Db 184 TTTGTAGTTCCTGCTGCCCATCATTTGGCATATGGCAGCTCAGCTCCTGAAATGGGAACG 243

Qy 81 LysAsnCysThrValGlySerValAsnAlaAspIleSerProSerProGluGlyLysGly 100
 Db 244 AAGAAATTCAGCGTTGGCTCAGTTAATGCAGATATATCTCCAGTCCGGAAGGCAAGGA 303

Qy 101 AsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsnMet 120
 Db 304 AATGGCAGTGAAGATGAATGAGATTTGGAAGAGCTGTGATGGAACCGCATGAGCAACATG 363

Qy 121 GluSerArgIleGlnTyrLeuSerAspAsnGluAlaAsnLeuLeuLeuLeuAsnPhe 140
 Db 364 GAAAGCAGAAATCCAGTATCTTTTCAGATTAATGAAGCCCAATCTCTAGATGCTTAAGATTC 423

Qy 141 GlnAsnPheSerIleThrThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsnSer 160

```
Db 424 CAAAATTTTCAGCAATAACAACCTGATCAAGAGATTTAATGATGTTCTTTTCCAGCTAAATTC 483
Qy 161 LeuLeuSerSerIleGlnGluHieGluAenIleIleGlyAspIleSerLysLeuVal 180
Db 484 TTACTTTCTCTCCATCCAGAGAAACATGAGAAATATCATAGGGGATATCTCAAGTCAATAGTA 543
Qy 181 GlyLeuAenThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAenGlyArgVal 200
Db 544 GGTCTGAACACCAACAGTACTTGTATTTGCAGTTTCAGATTGAAACACTGAATGGCAGAGTC 603
Qy 201 GlnGluAenAlaPheLysGlnGlnGluMetArgLysLeuGluGluAurGileTyAsn 220
Db 604 CAAGAGAATGCATTTAAACAACAAGAGAGATGCGTAAATTAGAGGCGGTATATACAAT 663
Qy 221 AlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyLeuGluGlnGluLys 240
Db 664 GCATCAGCAGAAATTAAGTCTCTAGATGAAACCAAGTATATTTGGAAACAGGAATAAAA 723
Qy 241 GlyGluMetLysLeuLeuAenAenIleThrAenAspLeuArgLeuLysAspTIPGluHis 260
Db 724 GGGGAATGAACCTGTTGAATATATCTAATGATCTGAGGCTGAAGGATTGGGAACAT 783
Qy 261 SerGlnThrLeuLysAenIleThrLeuLeuGlnGlyAla----- 273
Db 784 TCTCAGACATTTGAAAAATATCACTTACTCCAAGTCTCTCTGGACCTCCAGGTGAAAAA 843
Qy 274 -----ArgLysCysSerLeuThrGlyLysThrAenAspLeuGly---SerAsnMet 290
Db 844 GGAGATAGAGGCC-TCCTGGACA---AAATGGTATACCAAGCTTCCAGGCTTAATAGG 899
Qy 291 ThrIleGlyAlaValAenSerArgGlyGluPheThrGlyThrTyIleThrAlaValThr 310
Db 900 TACTCCAGTCTTAAAGTGATCGGGGATCTCTGGTTACTCGGATTCGAGATTCC 959
Qy 311 AlaThrSerAenGluIleLysGluSer-ProLeuHieGlyThrGlnAenThrIleAenLy 330
Db 960 AGGACCAATGGGAAGACCGGAAAGCCAGGACTTAATGGACAAAAAGGCCAGAGGCGA 1019
Qy 330 sArgThrGlnProThrPhe-----GlyPheThrVal-AsnTIPLysPheSerGlu 347
Db 1020 AAAACCCACTCCAAACCATCCAAACACATCTAATACAGTCCGATCGGTGGTGCGGCGG 1079
Qy 347 erThrThrValPheThrGlyGlnCysPheIleAspArgAsnGlyLysGluValLeuLys 367
Db 1080 CCCTCACAAGGCAGAGTGGAGATTTTCAGAAAGCCAGTGGGTCGCGTGTGACGA 1139
Qy 367 hrMetTIPLeuLeuArgSerSerVal-AsnAspIleGlyAsp---AspTIPLysAlaThr 385
Db 1140 CCGC-TGGGAACCTGCTGGAGACTGGTCTGCTCGCAGGAGCTTGGGATACAAAGTGTTC 1198
Qy 386 ArgValGlyIleAen-----IlePheThrArgLeuArgThrGln 398
Db 1199 AAAGTGTGCATAAGCGAGCTTATTTTGGAAAAAGGTACGGGTCCAA 1243

RESULT 4
AR365559 1807 bp DNA linear PAT 03-SBP-2003
LOCUS
DEFINITION Sequence 1 from patent US 5510466.
ACCESSION AR365559
VERSION AR365559.1 GI:34429330
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 1807)
AUTHORS Krieger, M. and Kodama, T.
TITLE Scavenger receptor protein and antibody thereto
JOURNAL Patent: US 5510466-A 1 23-APR-1996;
Massachusetts Institute of Technology; Cambridge, MA
FEATURES
source
1..1807
/organism="unknown"
```

```
ORIGIN
/mol_type="genomic DNA"
Alignment Scores:
Pred. No.: 3,79e-89 Length: 1807
Score: 1381.50 Matches: 288
Percent Similarity: 85.2% Conservative: 16
Best Local Similarity: 80.7% Mismatches: 34
Query Match: 67.1% Indels: 20
DB: 6 Gaps: 4

US-10-618-570-2 (1-400) x AR365559 (1-1807)
Qy 1 MetAlaGlnTIPAspAspPheProAspGlnGlnGluAspThrAspSerCysThrGluSer 20
Db 36 ATGGCACAGTGGGATGATCTTCTGATCAGCAAGAGAGCACTGCACAGTGTACAGAGTCT 95
Qy 21 ValLysPheAspAlaArgSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
Db 96 GTGAAGTTCGATGCTCGCTCAGTGACAGCTTGTCTCTCCCATCTCTAAAAATGGCCCA 155
Qy 41 ThrLeuGlnGluArgMetLysSerTyLysThrAlaLeuIleThrLeuTyLeuIleVal 60
Db 156 ACTCTTTCAGAGAGGATGAAAGTCTTATAAACTGCACCTGATCACCTTTATCTCATGTG 215
Qy 61 PheValValLeuValProIleIleGlyIleValAlaAlaGlnLeuLeuLysTIPGluThr 80
Db 216 TTTGTAGTCTCTGTCGCCATCATTTGGCATAGTGGCAGCTCAGCTCTGTAATGGGAAACG 275
Qy 81 LysAsnCysThrValGlySerValAsnAlaAspIleSerProSerProGluGlyLysGly 100
Db 276 AAGAATTCACCGTGGCTCAGTTAATGCAGATATATCTCCAAGTCCGGAAGGCAAGGA 335
Qy 101 AsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsnMet 120
Db 336 AATGGCAGTGAAGATGAATGAGATTTTCGAGAAGCTGTGATGGAAACCATGAGCAACATG 395
Qy 121 GluSerArgIleGlnTyLeuSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsnPhe 140
Db 396 GAAAGCAGATCCAGTATCTTTCAGATATGAGACCATCTCTAGATGCTNAGAAATTC 455
Qy 141 GlnAsnPheSerIleThrThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsnSer 160
Db 456 CAAAATTTTCAGCAATAACAACCTGATCAAGATTTAATGATGTTCTTTTCCAGCTAAATTC 515
Qy 161 LeuLeuSerSerIleGlnGluHieGluAenIleIleGlyAspIleSerLysLeuVal 180
Db 516 TTACTTTCTCTCCATCCAGGAACATGAGAAATATCATAGGGGATATCTCCAAGTCAATAGTA 575
Qy 181 GlyLeuAenThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAenGlyArgVal 200
Db 576 GGTCTGAACACCAACAGTACTTGTATTTGCAGTTTCAGTATTTGAAACACTGAATGGCAGAGTC 635
Qy 201 GlnGluAenAlaPheLysGlnGlnGluMetArgLysLeuGluGluAurGileTyAsn 220
Db 636 CAAGAGAATGCATTTAAACAACAAGAGAGATGCGTAAATTAGAGGCGGTATATACAAT 695
Qy 221 AlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyLeuGluGlnGluLys 240
Db 696 GCATCAGCAGAAATTAAGTCTCTAGATGAAACCAAGTATATTTGGAAACAGGAATAAAA 755
Qy 241 GlyGluMetLysLeuLeuAenAenIleThrAenAspLeuArgLeuLysAspTIPGluHis 260
Db 756 GGGGAATGAACCTGTTGAATATATCTAATGATCTGAGGCTGAAGGATTGGGAACAT 815
Qy 261 SerGlnThrLeuLysAenIleThrLeuLeuGlnGlyAla----- 273
Db 816 TCTCAGACATTTGAAAAATATCACTTTACTCCAAGTCTCTCTGGACCTCCAGGTGAAAAA 875
Qy 274 -----ArgLysCysSerLeuThrGlyLysThrAenAspLeuGly---SerAsnMet 290
Db 876 GGAGATAGAGGCC-TCCTGGACA---AAATGGTATACCAAGGCTTTCAGGCTTAATAGG 931
```

```

Qy 291 ThrileGlyAlaValAsnSerArgGlyCluPhethrGlyThrThrIleThrAlaValThr 310
Db 932 TACTCCAGGCTTAAAGGTGATCGGGGATCTCTGGTTTACCTGGAGTCCGAGATCCC 991

Qy 311 AlaThrSerAsnGluLeuLeuSerProLeuHisGlyThrGlnAsnThrIleAsnLys 330
Db 992 AGGACCATGGGAAGACCGGAGCCAGGACTTATTGGACAAAGGCCAGAGGGGAGA 1051

Qy 331 ArgThrGlnProThrPheGlyPheThrValAsnTrpLysPheSerGluSer 347
Db 1052 AAAAGGGAG-----TGGAGCATGCAAGACC 1078

RESULT 5
BTSCAVR1
LOCUS B.taurus mRNA 1592 bp mRNA linear MAM 18-APR-2005
DEFINITION B.taurus mRNA for macrophage scavenger receptor type I.
ACCESSION X51689 X54182
VERSION X51689.1 GI:734
KEYWORDS lipid metabolism; macrophage scavenger receptor type I; membrane glycoprotein.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
REFERENCE 1 (bases 1 to 1592)
AUTHORS Kodama,T.; Freeman,M., Rohrer,L., Zabrecky,J., Matsudaira,P. and Krieger,M.
TITLE Type I macrophage scavenger receptor contains alpha-helical and collagen-like coiled coils
JOURNAL Nature 343 (6258), 531-535 (1990)
PUBMED 2300204
REFERENCE 2 (bases 1 to 1592)
AUTHORS Krieger,M.
TITLE Direct Submission
JOURNAL Submitted (07-AUG-1990) Krieger M., Massachusetts Institute of Technology, Room E25-236, 77 Massachusetts Ave, Cambridge MA 02139, U S A

FEATURES
source
1..1592
location/Qualifiers
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="caxon:9913"
/clone="pBSR7"
/tissue type="lung"
/clone_lib="1ZAPII"
22..1383
/codon_start=1
/product="scavenger receptor type I"
/protein_id="CAA35987.1"
/db_xref="GI:735"
/db_xref="GOA:P21758"
/db_xref="InterPro:IPR001190"
/db_xref="InterPro:IPR003543"
/db_xref="InterPro:IPR008160"
/db_xref="InterPro:IPR008161"
/db_xref="UniProt/Swiss-Prot:P21758"
/tranlation="MAOWDFPDQEDTSCVESKFDARSVTALLPPHKNKPTLOE
RMKSYKTLALTYLVFVLPVPIIGIVAQLLKWETKNTCTGVSYNADISPSPEKGGNG
SEDMRFREAWGMSRMSRQYLSNEALLDAKQNFNSITTDQRFNDVLFQNS
LLSQIHEHNIIGDISLVGLNTVLDLQFSIETKRGVQENAFKQOENRKLBERI
YNASAEIKSLDEKQVLEQEIKGEMKLNNITNDLRKNDWEHSQTLKNIITLLQPPGP
PGEKGRDPPQNGIIPGPGIGITPGLKGDGRIQSLPGVGPFGPMGKTGKPLNGOK
GQKEKGSQMSQNSVRLVGGSGPHEGRVEIFHEGOWGTVCDDRLMRGLVVCVS
LYKGVQSVHRAFGKGTGTPIWLNEVFCFKESSISECRIRQGVGRACSHDEDAVGT
CTT"

CDS
3..68-89 Length: 1592
Pred. No.: 138.00 Matches: 272
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0

ORIGIN
Alignment Scores:
3..68-89 Length: 1592
Score: 138.00 Matches: 272
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0

```

```

Query Match: 67.1% Indels: 0
DB: 4 Gaps: 0
US-10-618-570-2 (1-400) x BTSCAVR1 (1-1592)

Qy 1 MetAlaGlnTrpAspAspPheProAspGlnGlnGluAspThrAspSerCysThrGluSer 20
Db 22 ATGGCAGAGTGGGATGACTTTCTCTGATCAGCAGAGGACACTGACAGCTGTACAGATCT 81

Qy 21 ValLysPheAspAlaArgSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
Db 82 GTGAAGTTGCGATGCTCGCTCAGTCAGCAGCTTTGCTTCTCTCCCATCTCTAAATGGCCCA 141

Qy 41 ThrLeuGlnGluArgMetLysSerTyrLysThrAlaLeuLeuThrLeuTyrLeuVal 60
Db 142 ACTCTTCAAGAGAGAGTGAAGTCTTATAAACTGCACCTGATCACCCTTTATCTCATTTG 201

Qy 61 PheValValLeuValProIleGlyIleValAlaAlaGlnLeuLeuLysTrpGluThr 80
Db 202 TTTGTAGTTCTCGTGGCCCATCATTTGGCATATGTGGCAGCTCAGCTCTCTGAAATGGGAAAC 261

Qy 81 LysAsnCysThrValGlySerValAsnAlaAspIleSerProSerProGluGlyLysGly 100
Db 262 AAGATTTCACGGTGGCTCAGTTAATGCAGATATATCTCCAGTCCGGAAGGCCAAAGGA 321

Qy 101 AsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsnMet 120
Db 322 AATGCCAGTGAAGATGAATGAGATTTTCGAGAAGCTGTGATGGAACGCATGAGCAACATG 381

Qy 121 GluSerArgIleGlnTyrLeuSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsnPhe 140
Db 382 GAAAGCAGATCCAGTATCTTTTCAGATAATGAAGCCCAATCTCTAGATGCTTAAGAAATTC 441

Qy 141 GlnAsnPheSerIleThrThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsnSer 160
Db 442 CAAATTTTCAGCATACACACTGATCAAGATTTTAATGATGTCTTTTCCAGCTAAATTC 501

Qy 161 LeuLeuSerSerIleGlnGluHisGluAsnIleLeuGlyAspIleSerLysSerLeuVal 180
Db 502 TTACTTTCTCCATCCAGGAACATGAGATATATCATGGGGATATCTCCAGCTATTAGTA 561

Qy 181 GlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArgVal 200
Db 562 GGTCTGAACACCCAGACTTGTATTCAGTTCAGTATTGAAACACTGAATGGCAGAGTTC 621

Qy 201 GlnGluAsnAlaPheLysGlnGlnGluMetArgLysLeuGluGluArgIleTyrAsn 220
Db 622 CAAGAGATGCATTTAAACAACAAGAGAGATGCTGAATTTAGAGGACGCTATATACAT 681

Qy 221 AlaSerAlaGluIleLysSerLeuAspLysGlnValTyrLeuGluGlnGluLeuLys 240
Db 682 GCATCAGCAGAAATTAAGTCTCTAGATGAAACCAAGTATATTTTGGAAACAGGAAATAAAA 741

Qy 241 GlyLysMetLysLeuLeuAsnAsnIleThrAsnAspLeuArgLeuLysAspTrpGluHis 260
Db 742 GGGGAAATGAACATGTTTGAATAATATCACTAATGATCTGAGGCTGAAGGATTTGGGAACAT 801

Qy 261 SerGlnThrLeuLysAsnIleThrLeuLeuGlnGly 272
Db 802 TCTCAGACATTTGAAATAATATCACTTTTCTCCAGGT 837

RESULT 6
BTSCAVR2
LOCUS B.taurus mRNA 2135 bp mRNA linear MAM 18-APR-2005
DEFINITION B.taurus mRNA for macrophage scavenger receptor, type II.
ACCESSION X54183
VERSION X54183.1 GI:736
KEYWORDS lipid metabolism; macrophage scavenger receptor type I; membrane glycoprotein.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
REFERENCE 1 (bases 1 to 2135)
AUTHORS Kodama,T.; Freeman,M., Rohrer,L., Zabrecky,J., Matsudaira,P. and Krieger,M.
TITLE Type II macrophage scavenger receptor contains alpha-helical and collagen-like coiled coils
JOURNAL Nature 343 (6258), 531-535 (1990)
PUBMED 2300204
REFERENCE 2 (bases 1 to 2135)
AUTHORS Krieger,M.
TITLE Direct Submission
JOURNAL Submitted (07-AUG-1990) Krieger M., Massachusetts Institute of Technology, Room E25-236, 77 Massachusetts Ave, Cambridge MA 02139, U S A

FEATURES
source
1..2135
location/Qualifiers
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="caxon:9913"
/clone="pBSR7"
/tissue type="lung"
/clone_lib="1ZAPII"
22..1383
/codon_start=1
/product="scavenger receptor type I"
/protein_id="CAA35987.1"
/db_xref="GI:735"
/db_xref="GOA:P21758"
/db_xref="InterPro:IPR001190"
/db_xref="InterPro:IPR003543"
/db_xref="InterPro:IPR008160"
/db_xref="InterPro:IPR008161"
/db_xref="UniProt/Swiss-Prot:P21758"
/tranlation="MAOWDFPDQEDTSCVESKFDARSVTALLPPHKNKPTLOE
RMKSYKTLALTYLVFVLPVPIIGIVAQLLKWETKNTCTGVSYNADISPSPEKGGNG
SEDMRFREAWGMSRMSRQYLSNEALLDAKQNFNSITTDQRFNDVLFQNS
LLSQIHEHNIIGDISLVGLNTVLDLQFSIETKRGVQENAFKQOENRKLBERI
YNASAEIKSLDEKQVLEQEIKGEMKLNNITNDLRKNDWEHSQTLKNIITLLQPPGP
PGEKGRDPPQNGIIPGPGIGITPGLKGDGRIQSLPGVGPFGPMGKTGKPLNGOK
GQKEKGSQMSQNSVRLVGGSGPHEGRVEIFHEGOWGTVCDDRLMRGLVVCVS
LYKGVQSVHRAFGKGTGTPIWLNEVFCFKESSISECRIRQGVGRACSHDEDAVGT
CTT"

```

| | | | | | |
|---|---|---|-----|---|-----|
| REFERENCE | Pecora; Bovidae; Bovinae; Bos. | Db | 396 | GAAGACGAAATCCAGTATCTTTTCAGATAATGAAGCCAAATCTCTAGTAGTCTAAGAATTTC | 455 |
| AUTHORS | 1 (bases 1 to 2135) | | | | |
| TITLE | Rohrer, L., Freeman, M., Kodama, T., Penman, M. and Krieger, M. | Qy | 141 | GLnSenPheSerIleThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsnSer | 160 |
| JOURNAL | Coiled-coil fibrous domains mediate ligand binding by macrophage | Db | 456 | CAAAATTTTCAGCAATAACAACACTGATCAAGAGATTTTAATGATGTCTTTTCCACGCTAAATTC | 515 |
| PUBMED | scavenger receptor type II | | | | |
| REFERENCE | Nature 343 (6258), 570-572 (1990) | Qy | 161 | LeuLeuSerSerIleGlnGluHisGluAsnIleIleGlyAspIleSerLysSerLeuVal | 180 |
| AUTHORS | 2 (bases 1 to 2135) | Db | 516 | TTACTTTCTCCATCCAGGACATGAGATATCATAGGGGATATCTCCAGTCATTAGTA | 575 |
| TITLE | Krieger, M. | | | | |
| JOURNAL | Direct Submission | Qy | 181 | GlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArgVal | 200 |
| REFERENCE | Submitted (07-AUG-1990) Krieger M., Massachusetts Institute of | Db | 576 | GGTCTGAACACCAACAGTACTTGATTGTCAGTTTCAGTATTGAAACACACACACAGAGTC | 635 |
| TITLE | Technology, Room E25-236, 77 Massachusetts Ave, Cambridge MA 02139, | | | | |
| JOURNAL | U S A | Qy | 201 | GlnGluAsnAlaPheLysGlnGlnGluMetArgLysLysGluGluArgIleTyrAsn | 220 |
| FEATURES | Location/Qualifiers | Db | 636 | CAAGAGAAATGCATTTAAACAACAAGAGAGAGATGCGTAAATTTAGAGGACGCTATATACAAT | 695 |
| source | 1..2135 | Qy | 221 | AlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyrLeuGluGlnGluIleLys | 240 |
| | /organism="Bos taurus" | Db | 696 | GCATCAGCAGAAATTAAGTCTCTAGATGAAACACAGTATATTTGGACAGGAATAAAA | 755 |
| | /mol_type="mRNA" | | | | |
| | /db_xref="taxon:9913" | Qy | 241 | GlyGluMetLysLeuLeuAsnAsnIleThrAsnAspLeuArgLeuLysAspTyrGluHis | 260 |
| | /clone="pBSR3, pBSR9" | Db | 756 | GGGGAATGAACATGTTGAATAATATATATCAATGATCTGAGGCTGAAGGATTGGAAACAT | 815 |
| | /tissue_type="lung" | Qy | 261 | SerGlnThrLeuLysAsnIleThrLeuLeuGlnGly | 272 |
| | /clone_fib="LZAPII" | Db | 816 | TCTCAGACATTTGAAATAATATCACTTTTACTCCAAGGT | 851 |
| | 36..1085 | | | | |
| CDS | /codon_start=1 | RESULT 7 | | | |
| | /product="bovine scavenger receptor type II" | AF037351 | | | |
| | /protein_id="CAA38108.1" | LOCUS | | | |
| | /db_xref="GI:737" | DEFINITION | | | |
| | /db_xref="GOA:P21758" | AF037351 | | | |
| | /db_xref="UniProt/Swiss-Prot:P21758" | ACCESION | | | |
| | /translation="MAQWDDPDDQEDTDSCTSVKFDARSVTALLPPHPKNGPTLQE | VERSION | | | |
| | RMKSKYKTALITLYLIVFVLVPIIGIVAAQLLKWETKNCTVGSVNADISPEKGGNG | AF037351 | | | |
| | SEDMERFREAVMERMSNMESRIQYLSNEANLLDAKNQFNSITTDQRFNDVLPQLNS | KEYWORDS | | | |
| | LLSSTOEHNITIGDISKSLGNTTVDLQFSIETLGRVOENAFKQOEERKLEERI | SOURCE | | | |
| | YNASAEIKSLDEKOVLYEQEIKGEMKLLNITNLRLLKOWEHSQTLKNITLLQPPGP | ORGANISM | | | |
| | PGEKGDPPQNGIPIPGFGLITGFLGLKGRGISGLPEVRGFPFGMGTGKPGUNGQK | Homo sapiens | | | |
| | GQKGEKSGSMQRPG" | Homo sapiens | | | |
| ORIGIN | | Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; | | | |
| Alignment Scores: | | Hominidae; Homo. | | | |
| Pre d. No.: | 4, 9e-89 | 1 (bases 1 to 1167) | | | |
| Score: | 1381.00 | Gough, P.J., Greaves, D.R. and Gordon, S. | | | |
| Percent Similarity: | 100.0% | A naturally occurring isoform of the human macrophage scavenger | | | |
| Best Local Similarity: | 100.0% | receptor (SR-A) gene generated by alternative splicing blocks | | | |
| Query Match: | 67.1% | modified LDL uptake | | | |
| DB: | 4 | J. Lipid Res. 39 (3), 531-543 (1998) | | | |
| | | 9548586 | | | |
| US-10-618-570-2 (1-400) x BTSCAVR2 (1-2135) | | 2 (bases 1 to 1167) | | | |
| Qy | 1 | Gough, P.J., Greaves, D.R. and Gordon, S. | | | |
| Db | 36 | Direct Submission | | | |
| Qy | 21 | Submitted (08-DEC-1997) Sir William Dunn School of Pathology, South | | | |
| Db | 96 | Parks Road, Oxford OX1 3RE, England | | | |
| Qy | 41 | Location/Qualifiers | | | |
| Db | 156 | 1..1167 | | | |
| Qy | 61 | /organism="Homo sapiens" | | | |
| Db | 216 | /mol_type="mRNA" | | | |
| Qy | 81 | /db_xref="taxon:9606" | | | |
| Db | 276 | /chromosome="8" | | | |
| Qy | 101 | /map="8p22" | | | |
| Db | 336 | /cell_line="THP-1; phorbol ester-treated 4 days" | | | |
| Qy | 121 | /cell_line="monocyte" | | | |
| | | 1..1167 | | | |
| | | /genes="SR-A" | | | |
| | | 1..1167 | | | |
| | | /genes="SR-A" | | | |
| | | /functions="blocks modified LDL uptake" | | | |
| | | /note="isoform; normal alternative splice product" | | | |
| | | /codon_start=1 | | | |
| | | /product="macrophage scavenger receptor type III" | | | |
| | | /protein_id="AAC09251.1" | | | |

```
/db_xref="GI:3004960"
/translation="MEQNDHFHQEDTDSCSESVKFDARSWTALLPPNKNPSLQE
KXKSFALIALYLLVFLIPLIGIAAQLLKWETKNCVSSVSNANDITQSLTGKN
DSBEEMFQEFVMEHSMKRIQHILDMEANLMDTFHFQNFMTDQRFNDILQLS
TLFSVOGHNAIDBISKSLISLNTLLDLQNTENLNGKIQENTFKQOEIILKEER
VYNSABTMAMKEOVHLEOBIKEVKVLNNITNDLRKOWEHSOTLNRNITLIOGPRG
PPGSGDRGPESGPRGPRGPIPGGLKGRGALPGPSRGLSGYAGRPNGSPKQ
KGEKSGNLTSTGFIWLNVEVFCRESSIESCKIQMGTRACSHSDAGVTCIL"
```

ORIGIN

```
Alignment Scores:
Pred. No.: 3,13e-65 Length: 1167
Score: 1042.50 Matches: 197
Percent Similarity: 87.2% Conservative: 41
Best Local Similarity: 72.2% Mismatches: 34
Query Match: 50.6% Indels: 1
DB: 8 Gaps: 1

US-10-618-570-2 (1-400) x AF037351 (1-1167)

Qy 1 MetAlaGlnTrpAspPheProaspGlnGlnAspThrAspSerCysThrGluSer 20
Db 1 ATGGAGCAGTGGGATCACTTTTCAATCAACAGGAGGACACTAGTAGCTCGCAATCT 60
Qy 21 VallysPheAspAlaArgSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
Db 61 GTGAAATTTGATGCTCGCTCAATGACAGCTTTGCTTCTCGAATCTCTAAACAGCCCT 120
Qy 41 ThrLeuGlnGluArgMetLysSerTyrlsThrAlaLeuLeuThrLeuTyrlsLeuVal 60
Db 121 TCCCTTCAGAGAACTGAATGCTTCAAGCTGCAGCTGATGCTTACCTCTCTCGTG 180
Qy 61 PheValValLeuValProIleIleValAlaAlaGlnLeuLeuLysTrpGluThr 80
Db 181 TTTCAGTTCATCTCCTCTCATGATGATGAGCTCAACTCTCGAATCTCTAAACAGCCCT 240
Qy 81 LysAsnCysThrValGlySerValAsnAla---AspIleSerProSerProGluGlyLys 99
Db 241 AAGAATTCCTCAGTTAGTTCAACTAATGCAATGATATATACTCAAGTCTCACGGGAAA 300
Qy 100 GlyAsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsn 119
Db 301 GGAATGACAGGAGGAAGAAATGAGATTTCAAGAAAGTCTTTATGGAACACATGAGCAAC 360
Qy 120 MetGluSerArgIleGlnTyrlsSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsn 139
Db 361 ATGGAGAGAGATCCAGCATATTTAGACATGAGGCACTCATGACACAGAGCAT 420
Qy 140 PheGlnAsnPheSerIleThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsn 159
Db 421 TTCCAAAATTTTCAGCATGACAACTGATCAAGATTTAATGACATTTCTTCGAGCTAAGT 480
Qy 160 SerLeuLeuSerSerIleGlnGluHisGluAsnIleIleGlyAspIleSerLysSerIleu 179
Db 481 ACCTTGTTTCTCAGTCAGGAGCATGGGAATGCAATAGATGAATCTCCAAGTCTCTTA 540
Qy 180 ValGlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArg 199
Db 541 ATAAGTTGATACCACTTCCTGATTTGATGCTCAACATAGAAATCTGAATGGCAAA 600
Qy 200 ValGlnGluAsnAlaPheLysGlnGlnGluMetArgLysLeuGluGluArgIleTyrl 219
Db 601 ATCCAAGAGATACCTTCAACCAACAGAGGAAATCAGTAAATTTAGAGGAGCGTGTTCAC 660
Qy 220 AsnAlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyrlsLeuGluGluIle 239
Db 661 AATGTATCAGCAGAAATTTATGGCTATGAAAGAAAGCAAGTGCATTTTGGACAGGAATA 720
Qy 240 LysGlyGluMetLysLeuAsnAsnIleThrAsnAspLeuArgLeuLysAspTrpGlu 259
Db 721 AAGAGAGAGTGAAGTACTGATTAACATCACTAATGATCTCAGACTGAAGATTTGGAA 780
Qy 260 HisSerGlnThrLeuLysAsnIleThrLeuLeuGlnGly 272
```

```
Db 781 CATCTCAGACCTTGAGAAATATCACTTTAATCAAGGT 819
```

RESULT 8

```
E03268
LOCUS E03268 1347 bp RNA linear PAT 29-SEP-1997
DEFINITION cDNA sequence coding for human scavenger receptor, type II.
ACCESSION E03268
VERSION E03268.1 GI:2171485
KEYWORDS JP 1991290184-A/2.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.

REFERENCE

1 (bases 1 to 1347)
Matsumoto, A. and Kodama, T.
CELL OF SCAVENGER RECEPTOR-PRODUCING ANIMAL
Patent: JP 1991290184-A 2 19-DEC-1991;

JOURNAL

CHUGAI PHARMACEUT CO LTD
OS Homo sapiens (human)
PN JP 1991290184-A/2
PD 19-DEC-1991

COMMENT

PF 06-APR-1990 JP 1990090274
PI MATSUMOTO AKIYO, KODAMA TATSUHIKO
PC C12N5/10, C12N15/12, C12Q1/06;
CC strandedness: Single;
CC topology: Linear;

FEATURES

CC hypothetical: No;
CC anti-sense: No;
CC *source: cell_type=monocyte leukemic cells;
CC *source: cell_line=THP-1;
FH Key Location/Qualifiers
FT 5'UTR 1..46
FT CDS 47..1123
FT 3'UTR 1124..1301.
FT /product='human scavenger receptor, type II' FT

source

1..1347 Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic RNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.: 3.64e-65 Length: 1347
Score: 1042.50 Matches: 197
Percent Similarity: 87.2% Conservative: 41
Best Local Similarity: 72.2% Mismatches: 34
Query Match: 50.6% Indels: 1
DB: 6 Gaps: 1

US-10-618-570-2 (1-400) x E03268 (1-1347)

```
Qy 1 MetAlaGlnTrpAspPheProaspGlnGlnAspThrAspSerCysThrGluSer 20
Db 47 ATGGAGCAGTGGGATCACTTTTCAATCAACAGGAGGACACTAGTAGCTCGCAATCT 106
Qy 21 VallysPheAspAlaArgSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
Db 107 GTGAAATTTGATGCTCGCTCAATGACAGCTTTGCTTCTCGAATCTCTAAACAGCCCT 166
Qy 41 ThrLeuGlnGluArgMetLysSerTyrlsThrAlaLeuLeuThrLeuTyrlsLeuVal 60
Db 167 TCCCTTCAGAGAACTGAATGCTTCAAGCTGCAGCTGATGCTTACCTCTCTCGTG 226
Qy 61 PheValValLeuValProIleIleValAlaAlaGlnLeuLeuLysTrpGluThr 80
Db 227 TTTCAGTTCATCTCCTCTCATGATGATGAGTGTGACACTCACTCTCTGAGTGGGAAACG 286
Qy 81 LysAsnCysThrValGlySerValAsnAla---AspIleSerProSerProGluGlyLys 99
```



```

DEFINITION Antihuman scavenger receptor antibody.
ACCESSION BD000149.1 GI:18623228
VERSION JP 2000312595-A/2
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
REFERENCE 1 (bases 1 to 1347)
AUTHORS Kodama,T., Matsumoto,A. and Suzuki,H.
TITLE Antihuman scavenger receptor antibody
JOURNAL Patent: JP 2000312595-A 2 14-NOV-2000;
CHUGAI PHARM CO LTD
OS Homo sapiens (human)
PD JP 2000312595-A/2
PE 14-NOV-2000
PF 14-APR-2000 JP 2000114233
PI
PR
PT TATSUHIKO KODAMA, AKIYO MATSUMOTO, HIROSHI SUZUKI PC
C12N15/09, C07K16/28, C12N5/10, C12P21/08// (C12N5/10, C12R1.91), PC
(C12P21/08, C12R1.91), C12N15/00, C12N5/00, (C12N5/00, C12R1.91) CC
FH Key Location/Qualifiers
FT source 1..1301
FT /organism="Homo sapiens (human)".

FEATURES
source
LOCATION/Qualifiers
1..1347
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 3.64e-65 Length: 1347
Score: 1042.50 Matches: 197
Percent Similarity: 87.2% Conservatives: 41
Best Local Similarity: 72.2% Mismatches: 34
Query Match: 50.6% Indels: 1
Gaps: 6

US-10-618-570-2 (1-400) x BD000149 (1-1347)

Qy 1 MetAlaGlnTTPAspAepPheProAspGlnGlnGluAspThrAspSerCysThrGluSer 20
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
47 ATGGAGCAGTGGGATCACTTTCAACAATCAACAGGAGGACACTGATAGTCTCCGAATCT 106
Qy 21 ValLysPheAspAlaAArgSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
107 GTGAAATTGATGCTCGCTCAATGACAGCTTTGCTTCTCCGAATCTCTAAACACAGCCCT 166
Qy 41 ThrLeuGlnGluArgMetLysSerTyrLysThrAlaLeuLeuThrLeuTyrLeuVal 60
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
167 TCCCTTCAAGAGAATCACTGCTTCAAAAGCTGCACTGATGCTTACCTCTCTCGTG 226
Qy 61 PheValValLeuValProIleGlyLeValAlaAlaGlnLeuLeuLysTyrGluThr 80
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
227 TTTCAGTCTCATCTCCCTCTATTGGAATAGTGGCAGCTCAACTCTCGAAGTGGGAAACG 286
Qy 81 LysAsnCysThrValGlySerValAsnAla---AspIleSerProSerProGluGlyLys 99
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
287 AAGAAATGCTCAGTTAGTTAGTCACTAATCAATGATATAACTCAAAAGTCTCACCGGAAA 346
Qy 100 GlyAsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsn 119
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
347 GGAATGACACGAGAGAGAAATGAGATTCAAGAGTCTTTATGGAACACATGAGCAAC 406
Qy 120 MetGluSerArgIleGlnTyrLeuSerAspAsnGlnAlaAsnLeuLeuAspAlaLysAsn 139
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
407 ATGGAGAGAGAAATCCACATATTTTAGACATGGAAGCCAACCTCATGGACACAGAGCAT 466
Qy 140 PheGlnAsnPheSerIleThrThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsn 159
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
467 TTCCAAATTTTCAGCATGACAACTGATCAAAAGATTTATGACATTTCTTCGAGTAAAGT 526

```

```

Qy 160 SerLeuLeuSerSerIleGlnGlnHISGluAsnIleIleGlyAspIleSerLysSerLeu 179
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
527 ACCTTTGTTTCTCAGTCCAGGACATGGGAATCAATAGATGAATCTCCAAGTCTCTTA 586
Qy 180 ValGlyLeuLeuThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArg 199
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
587 ATAAGTTTGAATACCAATTCCTGATTCGAGTCAACATAGAAATCTGAAATGGCAAA 646
Qy 200 ValGlnGluAsnAlaPheLysGlnGlnGluMetArgLysLeuGluGluArgIleTyr 219
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
647 ATCCAAGAGATACCTTTCAACACACAGAGAAATCAGTAAATTAGAGGCGGTGTTTAC 706
Qy 220 AsnAlaSerAlaGluIleLysSerLeuAspGlnLysGlnValTyrLeuGluGlnGluIle 239
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
707 AATGTATCAGCAGAAATTTATGGCTATGAAAGAGAACAGAGTGCATTTTGGAAACAGAAATA 766
Qy 240 LysGlyGluMetLysLeuLeuAsnIleThrAsnAspLeuArgLysLeuAspTyrGlu 259
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
767 AAGGAGAGTGAAGTACTGAAATCAATCACTAATGATCTCAGACTGAAAGATTGGGAA 826
Qy 260 HisSerGlnThrLeuLysAsnIleThrLeuLeuGlnGly 272
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
827 CATCTCAGACCTTGAGAAATATCACTTTAATTCAGGT 865

```

```

RESULT 11
CQ722728 LOCUS CQ722728 1367 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 8662 from Patent WO02068579.
ACCESSION CQ722728
VERSION CQ722728.1 GI:42283585
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
REFERENCE 1
AUTHORS Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
TITLE Kits, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof
JOURNAL Patent: WO 02068579-A 8662 06-SEP-2002;
PE Corporation (NY) (US)
FEATURES
source
1..1367
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

```

```

ORIGIN
Alignment Scores:
Pred. No.: 3.69e-65 Length: 1367
Score: 1042.50 Matches: 197
Percent Similarity: 87.2% Conservatives: 41
Best Local Similarity: 72.2% Mismatches: 34
Query Match: 50.6% Indels: 1
Gaps: 6
DB:
US-10-618-570-2 (1-400) x CQ722728 (1-1367)

```

```

Qy 1 MetAlaGlnTTPAspAepPheProAspGlnGlnGluAspThrAspSerCysThrGluSer 20
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
67 ATGGAGCAGTGGGATCACTTTCAACAATCAACAGGAGGACACTGATAGTCTCCGAATCT 126
Qy 21 ValLysPheAspAlaAArgSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
127 GTGAAATTGATGCTCGCTCAATGACAGCTTTGCTTCTCCGAATCTCTAAACACAGCCCT 186
Qy 41 ThrLeuGlnGluArgMetLysSerTyrLysThrAlaLeuLeuThrLeuTyrLeuVal 60
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
187 TCCCTTCAAGAGAATCAAGTCTCTCAAAAGCTGCACTGATGCTTACCTCTCTCGTG 246

```

Qy 61 PheValValLeuValProIleGlyIleValAlaAlaGlnLeuLeuLeuTyrGluThr 80
Db 247 TTTGAGTTCTCATCCCTCTCATTTGGAATAGTGGCAGCTCACTCTCTGAAAGTGGGAAACG 306
Qy 81 LysAsnCysThrValGlySerValAsnAla---AspIleSerProSerProGluGlyLys 99
Db 307 AAGAATTCTCAGTTAGTTCACTAATGCAATGATATAAATCAAGTCTCAGCGGAAAC 366
Qy 100 GlyAsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsn 119
Db 367 GGAATGACAGCGAAGAGGAATGAGATTTCAAGAAGTCTTTATGGAAACACATCAGCAAC 426
Qy 120 MetGluSerArgIleGlnTyrLeuSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsn 139
Db 427 ATGGAGAAGAGATCCAGCATATTTAGCATGTGGAGCCCACTCATGGACACAGAGCAT 486
Qy 140 PheGlnAsnPheSerIleThrThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsn 159
Db 487 TTCCAAATTTTCAGCATGACAACTGATCAAAGATTTTAATGACATTTCTTCGACGTAAGT 546
Qy 160 SerLeuLeuSerSerIleGlnGluHisGluAsnIleIleGlyAspIleSerLysSerLeu 179
Db 547 ACCTTGTTTCTCAGTCCAGGACATGGGAATGCAATAGATGAATCTCCAAGTCTCTTA 606
Qy 180 ValGlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArg 199
Db 607 ATAAGTTGATATCCACATCTGCTGATTTGGAGTCAACATAGAAATCTGAATGGCAAC 666
Qy 200 ValGlnGluAsnAlaPheLysGlnGlnGluMetArgLysLeuGluGluArgIleTyr 219
Db 667 ATCCAAGAGATACCTTCAACCAACAGAGGAAATCAGTAAATTTAGAGGCGGTGTAC 726
Qy 220 AsnAlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyrLeuGluGlnGluIle 239
Db 727 AATGTATCAGCAGAAATTTATGCTATGAAGAAGCAAGCAATGCTATGATCTCAGACTGAAAGATTGGAA 846
Qy 260 HisSerGlnThrLeuLysAsnIleThrLeuLeuGlnGly 272
Db 847 CATCTCAGACCTTGAGAAATATCACTTTAATTCAGGT 885

RESULT 12
172357
LOCUS I72357 1367 bp DNA linear PAT 03-APR-1998
DEFINITION Sequence 3 from patent US 5683903.
ACCESSION I72357
VERSION I72357.1 GI:3008496
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.

REFERENCE
1 (bases 1 to 1367)
Lysko,P.G., Elshourbagy,N.A. and Brawner,M.E.

TITLE
Attachment enhanced 293 cells

JOURNAL
Patent: US 5683903-A 3 04-NOV-1997;

FEATURES
Location/Qualifiers
1..1367

source
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Alignment Scores:
Pred. No.: 3,69e-65 Length: 1367
Score: 1042.50 Matches: 197
Percent Similarity: 87.2% Conservative: 41
Best Local Similarity: 72.2% Mismatches: 34
Query Match: 50.6% Indels: 1
DB: 6 Gaps: 1

US-10-618-570-2 (1-400) x I72357 (1-1367)

Qy 1 MetAlaGlnTrpAspPheProAspGlnGlnGluAspThrAspSerCysThrGluSer 20
Db 67 ATGAGCAGTGGGATCACTTTCAATCAACAGAGGACACTGATAGCTGCTCGAATCT 126
Qy 21 ValLysPheAspAlaArgSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
Db 127 GTGAAATTTGATGCTCGCTCAATGACAGCTTTGCTTCTCGAATCTCTAAACACAGCCT 186
Qy 41 ThrLeuGlnGluArgMetLysSerTyrLysThrAlaLeuIleThrLeuTyrLeuIleVal 60
Db 187 TCCCTTCAAGAGAACTGAAGTCTCTCAAAAGCTGCACCTGATTTGCCCTTTTACTCTCTCGTG 246
Qy 61 PheValValLeuValProIleIleGlyIleValAlaAlaGlnLeuLeuLeuTyrGluThr 80
Db 247 TTTGCACTTCTCATCCCTCTCATTTGGAATAGTGGCAGCTCACTCTCTGAAGTGGAAACG 306
Qy 81 LysAsnCysThrValGlySerValAsnAla---AspIleSerProSerProGluGlyLys 99
Db 307 AAGAATTTCTCAGTTAGTTCACTAATGCAATGATATAAATCAAGTCTCAGCGGAAAC 366
Qy 100 GlyAsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsn 119
Db 367 GGAATGACAGCGAAGAGGAATGAGATTTCAAGAAGTCTTTATGGAAACACATCAGCAAC 426
Qy 120 MetGluSerArgIleGlnTyrLeuSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsn 139
Db 427 ATGGAGAAGAGATCCAGCATATTTAGCATGTGGAGCCCACTCATGGACACAGAGCAT 486
Qy 140 PheGlnAsnPheSerIleThrThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsn 159
Db 487 TTCCAAATTTTCAGCATGACAACTGATCAAAGATTTTAATGACATTTCTTCGACGTAAGT 546
Qy 160 SerLeuLeuSerSerIleGlnGluHisGluAsnIleIleGlyAspIleSerLysSerLeu 179
Db 547 ACCTTGTTTCTCAGTCCAGGACATGGGAATGCAATAGATGAATCTCCAAGTCTCTTA 606
Qy 180 ValGlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArg 199
Db 607 ATAAGTTTGAATACCACTTGTGATTTGAGTCAACATAGAAATTTAGAGGCGGTGTAC 666
Qy 200 ValGlnGluAsnAlaPheLysGlnGlnGluMetArgLysLeuGluGluArgIleTyr 219
Db 667 ATCCAGAGATACCTTCAACCAACAGAGGAAATCAGTAAATTTAGAGGCGGTGTAC 726
Qy 220 AsnAlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyrLeuGluGlnGluIle 239
Db 727 AATGTATCAGCAGAAATTTATGCTATGAAGAAGCAACAGTGCATTTTGGAAACAGAAATA 786
Qy 240 LysGlyGluMetLysLeuLeuAsnIleThrAsnAspLeuArgLeuLysAspTyrGlu 259
Db 787 AAAGGAGAGTGAAGTACTGAAATACATCACTAATGATCTCAGACTGAAAGATTGGAA 846
Qy 260 HisSerGlnThrLeuLysAsnIleThrLeuLeuGlnGly 272
Db 847 CATCTCAGACCTTGAGAAATATCACTTTAATTCAGGT 885

RESULT 13

AR270697

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

AR270697 1367 bp DNA linear PAT 10-APR-2003
Sequence 1260 from patent US 6500938.

AR270697.1 GI:29701931

Unknown.

Unknown.

Unclassified.

1 (bases 1 to 1367)

Au-Young,J. and Seilhamer,J.J.

Composition for the detection of signaling pathway gene expression

Patent: US 6500938-A 1260 31-DEC-2002;

Incyte Genomics, Inc.; Palo Alto, CA;
WOX;

```

FEATURES
  source
    Location/Qualifiers
      1..1367
        /organism="unknown"
        /mol_type="genomic DNA"

ORIGIN
  Alignment Scores:
    Pred. No.: 3.69e-65 Length: 1367
    Score: 1042.50 Matches: 197
    Percent Similarity: 87.2% Conservatives: 41
    Best Local Similarity: 72.2% Mismatches: 34
    Query Match: 50.6% Indels: 1
    DB: 6 Gaps: 1

US-10-618-570-2 (1-400) x AR270697 (1-1367)

Qy 1 MetAlaGlnTrpAspAspPheProaspGlnGlnAspThrAspSerCysThrGluSer 20
Db 67 ATGGAGCAGTGGGATCATTTCACATCAACAGGAGGACATGATAGCTCCGAACTCT 126
Qy 21 VallysPheAspAlaAargSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
Db 127 GTGAAATTTGATCGCTCAATGACAGCTTTGCTCTCCGAACTCTTAAACACAGCCCT 186
Qy 41 ThrLeuGlnGluArgMetLysSerTyrLysThrAlaLeuLeuThrLeuTyrLeuVal 60
Db 187 TCCCTTCAAGAGAAACTGAACTGCTTCAAGCTGCACTGATTCGCTTTACCTCTCGTG 246
Qy 61 PheValValLeuValProIleleGlyLeValAlaAlaGlnLeuLeuValTyrGluThr 80
Db 247 TTTCAGTTCATCCCTCTCATTTGGAATAGTGGCAGCTCAACTCTGAAAGTGGAAACG 306
Qy 81 LysAsnCysThrValGlySerValAsnAla---AspIleSerProSerProGluGlyLys 99
Db 307 AAGATTGCTCAGTTAGTTCAACTAATGCAATGATATACTCAAGTCTCAGCGGAAA 366
Qy 100 GlyAsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsn 119
Db 367 GGAATGACAGGAGAGAAATGAGATTTCAAGAAAGTCTTTATGGAACACATGAGCAAC 426
Qy 120 MetGluSerArgIleGlnTyrLeuSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsn 139
Db 427 ATGGAAGAGAAATCCAGCATATTTTACATGATGAAGCAACCTCATGGACACAGAGCAT 486
Qy 140 PheGlnAsnPheSerIleThrThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsn 159
Db 487 TTCCAAATTTTCAGCATGACACTGATCAAGATTTAATGACATCTCTCGACCTAAGT 546
Qy 160 SerLeuLeuSerSerIleGlnGluHisGluAsnIleleGlyAspIleSerLysSerLeu 179
Db 547 ACCTTGTCTTCTCAGTCCAGGACATGGGAATGCAATAGATGAATCTCCAAAGTCTTA 606
Qy 180 ValGlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArg 199
Db 607 ATAAAGTTGAATACCATCTCTGATTTGAGTTCAGCTCAACATAGAAAATCTGAATGGCAA 666
Qy 200 ValGlnGluAsnAlaPheLysGlnGlnGluMetArgLysLeuGluGluArgIleTyr 219
Db 667 ATCCAGAGATACCTTCAACACACAGAGAAATCAGTAATTAGAGAGCGGTTCAC 726
Qy 220 AsnAlaSerAlaGluLeuLysSerLeuAspGluLysGlnValTyrLeuGluGlnGluLe 239
Db 727 AATGATATCAGCAGAAATTTATGCTATGAAAGAAAGCAAGTGCAATTTGGAACAGGAATA 786
Qy 240 LysGlyGluMetLysLeuLeuAsnAsnIleThrAsnAspLeuArgLeuLysAsnTyrGlu 259
Db 787 AARGAGAGAGTGAAGTACTGAATAACATCATCTAATGATCTCAGACTGAAGATGGGAA 846
Qy 260 HisSerGlnThrLeuLysAsnIleThrLeuLeuGlnGly 272
Db 847 CATCTCAGACCTTGAGAAATATATCATCTTAATTCAAGT 885

RESULT 14

```

HUMPHSR2

LOCUS Homo sapiens phsR2 mRNA for scavenger receptor type II, complete cds.
 1367 bp mRNA linear PRI 18-DEC-2002

ACCESSION D90188
 VERSION D90188.1 GI:219991

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1367)

AUTHORS Matsumoto, A., Naito, M., Itakura, H., Ikemoto, S., Asooka, H., Hayakawa, I., Kanamori, H., Aburatani, H., Takaku, F., Suzuki, H., Kobari, Y., Miyai, T., Takahashi, K., Cohen, H. E., Wydro, R., Housman, E. D. and Kodama, T.

TITLE Human macrophage scavenger receptors: primary structure, expression, and localization in atherosclerotic lesions

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 87 (23), 9133-9137 (1990)

PUBMED 2251254

COMMENT These data kindly submitted in computer readable form by: Akiyo Matsumoto

FEATURES source

Location/Qualifiers

1-23-1 Toyama

Shinjuku-ku, Tokyo 162

Phone: 81-3-3203-5725

Fax: 81-3-3207-3520

Location/Qualifiers

1..1367

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/note="monocytic cell line THP-1 (4 days after phorbol ester treatment)"

gene 1..1367

CDS 67..1143

/gene="phsR2"

/gene="phsR2"

/codon_start=1

/product="scavenger receptor type II"

/protein_id="BAAL4209.1"

/db_xref="GI:219992"

/translation="MEQWDFHNOQEDTDCSESVKFDARSMATLLPPNPKNSLOE
 KLSFKAALIALYLLPAVLPIIGIIVAOLKWKETKNCVSSNTANDITQSLTKGN
 DSEERKPFQVFMHSMNEKRIQHLIDMEANLMDTEHFQNFSTTDQRENDILLQLS
 TLFSVGHGNAIDEISLSLNTLLDLQNLNENLNGIKIQENTFKQKEISKLSE
 VYNVSAIEMAKBEQVHLEQIEKGEVKNNTNDLRKDWHSQTLRNTLIQPPG
 PPGKGDGRGTGSGPRGPGPIGPPGLKGDRAIGPPGSRGLPGYAGRPNGSPKGO
 KGEKSGNTLRPVQLTDHIRAGPS"

ORIGIN

Alignment Scores:

Pred. No.: 3.69e-65 Length: 1367

Score: 1042.50 Matches: 197

Percent Similarity: 87.2% Conservatives: 41

Best Local Similarity: 72.2% Mismatches: 34

Query Match: 50.6% Indels: 1

DB: 8 Gaps: 1

US-10-618-570-2 (1-400) x HUMPHSR2 (1-1367)

Qy 1 MetAlaGlnTrpAspAspPheProaspGlnGlnAspThrAspSerCysThrGluSer 20

Db 67 ATGGAGCAGTGGGATCATTTCACATCAACAGGAGGACACTGATAGCTCCGAACTCT 126

Qy 21 VallysPheAspAlaAargSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40

Db 127 GTGAAATTTGATCGCTCAATGACAGCTTTGCTCTCCGAACTCTTAAACACAGCCCT 186

Qy 41 ThrLeuGlnGluArgMetLysSerTyrLysThrAlaLeuLeuThrLeuTyrLeuVal 60

Db 187 TCCCTTCAAGAGAAACTGAACTGCTTCAAGCTGCACTGATTCGCTTTACCTCTCGTG 246

Qy 61 PheValValLeuValProIleleGlyLeValAlaAlaGlnLeuLeuValTyrGluThr 80

Db 247 TTTCAGTTCATCCCTCTCATTTGGAATAGTGGCAGCTCAACTCTGAAAGTGGAAACG 306

Qy 81 LysAsnCysThrValGlySerValAsnAla---AspIleSerProSerProGluGlyLys 99

Db 307 AAGATTGCTCAGTTAGTTCAACTAATGCAATGATATACTCAAGTCTCAGCGGAAA 366

Qy 100 GlyAsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsn 119

Db 367 GGAATGACAGGAGAGAAATGAGATTTCAAGAAAGTCTTTATGGAACACATGAGCAAC 426

Qy 120 MetGluSerArgIleGlnTyrLeuSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsn 139

Db 427 ATGGAAGAGAAATCCAGCATATTTTACATGATGAAGCAACCTCATGGACACAGAGCAT 486

Qy 140 PheGlnAsnPheSerIleThrThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsn 159

Db 487 TTCCAAATTTTCAGCATGACACTGATCAAGATTTAATGACATCTCTCGACCTAAGT 546

Qy 160 SerLeuLeuSerSerIleGlnGluHisGluAsnIleleGlyAspIleSerLysSerLeu 179

Db 547 ACCTTGTCTTCTCAGTCCAGGACATGGGAATGCAATAGATGAATCTCCAAAGTCTTA 606

Qy 180 ValGlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArg 199

Db 607 ATAAAGTTGAATACCATCTCTGATTTGAGTTCAGCTCAACATAGAAAATCTGAATGGCAA 666

Qy 200 ValGlnGluAsnAlaPheLysGlnGlnGluMetArgLysLeuGluGluArgIleTyr 219

Db 667 ATCCAGAGATACCTTCAACACACAGAGAAATCAGTAATTAGAGAGCGGTTCAC 726

Qy 220 AsnAlaSerAlaGluLeuLysSerLeuAspGluLysGlnValTyrLeuGluGlnGluLe 239

Db 727 AATGATATCAGCAGAAATTTATGCTATGAAAGAAAGCAAGTGCAATTTGGAACAGGAATA 786

Qy 240 LysGlyGluMetLysLeuLeuAsnAsnIleThrAsnAspLeuArgLeuLysAsnTyrGlu 259

Db 787 AARGAGAGTGAAGTACTGAATAACATCATCTAATGATCTCAGACTGAAGATGGGAA 846

Qy 260 HisSerGlnThrLeuLysAsnIleThrLeuLeuGlnGly 272

Db 847 CATCTCAGACCTTGAGAAATATATCATCTTAATTCAAGT 885

```

Db 187 TCCCTTCAAGAGAACTGAAGTCCTTCAAGCTGCACCTGATTGGCCCTTTACCTCTCGTG 246
Qy 61 PheValValLeuValProIleGlyIleValAlaAlaGlnLeuLeuValTrpGluThr 80
Db 247 TTTGCAGTTCATCTCCCTCTCATTTGGAATAGTGGCAGCTCAACTCTCTGAAGTGGGAACG 306
Qy 81 LysAsnCysThrValGlySerValAsnAla---AspIleSerProSerProGluGlyLys 99
Db 307 AAGAATTGCTCAGTTAGTTCAACTAATGCAATGATATATACTCAAGATCTCACGGGAAA 366
Qy 100 GlyAsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsn 119
Db 367 GGAATGACAGCGAGAGAAATGAGATTTTCAAGAAGTCTTTATGGAACACATGAGCAAC 426
Qy 120 MetGluSerArgIleGlnTyrLeuSerAspAsnGlnAlaAsnLeuLeuAspAlaLysAsn 139
Db 427 ATGGAAGAAGATCCACATATTTAGCATGTGAAGCCAACTCATGGACACAGAGCAT 486
Qy 140 PheGlnAsnPheSerIleThrThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsn 159
Db 487 TTCCAAAATTTTCAGCATGACAACTGATCAAGATTTAATGACATTTCTTCGACGTAAGT 546
Qy 160 SerLeuLeuSerSerIleGlnGluHisGluAsnIleIleGlyAspIleSerLysSerLeu 179
Db 547 ACCTTTGTTTCTCAGTCCAGGCACATGGGAATGCAATAGATGAATCTCCAAGTCTCTTA 606
Qy 180 ValGlnLeuAsnAlaPheLysGlnGlnGluMetArgLysLeuGluGluArgIleTyr 219
Db 607 ATAAGTTGAATACCATCTGCTGATTGTTGAGCTCAACATAGAAATCTGAATGGCAAA 666
Qy 200 ValGlnGluAsnAlaPheLysGlnGlnGluMetArgLysLeuGluGluArgIleTyr 219
Db 667 ATCCAAGAGATACCTTCAACACACAGAGGAATCAGTAATTTAGAGGCGTGTTCAC 726
Qy 220 AsnAlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyrLeuGluGlnGluIle 239
Db 727 AATGATATCAGCAGAAATTTATGGCTATGAAAGAAAGAACAAAGTGCAATTTGGAAACAGGAATA 786
Qy 240 LysGlyGluMetLysLeuLeuAsnAsnIleThrAsnAspLeuArgLeuLysAspTrpGlu 259
Db 787 AAGGAGAAGTGAAGATGTAAGTAACATCACTAATGATCTTCGACTGAAAGATTGGGAA 846
Qy 260 HisSerGlnThrLeuLysAsnIleThrLeuLeuGlnGly 272
Db 847 CATCTCAGACCTTGAGAATATCACITTAATTCAGGT 885

RESULT 15
E03267 LOCUS E03267 2028 bp RNA linear PAT 29-SEP-1997
DEFINITION cDNA sequence coding for human scavenger receptor, type 1.
ACCESSION E03267
VERSION E03267.1 GI:2171484
KEYWORDS JP 1991290184-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 2028)
Matsumoto,A. and Kodama,T.
CELL OF SCAVENGER RECEPTOR-PRODUCING ANIMAL
Patent: JP 1991290184-A 1 19-DEC-1991;
CHUGAI PHARMACEUT CO LTD
OS Homo sapiens (human)
PN JP 1991290184-A/1
PD 19-DEC-1991
PF 06-APR-1990 JP 1990090274
PI MATSUMOTO AKIYO, KODAMA TATSUHIKO
PC C12N5/10,C12N15/12,C12Q1/06;
CC strandedness: Single;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;

```

```

CC *source: cell type=monocyte leukemic cells;
CC *source: cell_line=THP-1;
FH Key Location/Qualifiers
FH 5'UTR 1..46
FH CDS 47..1402
FT /product='human scavenger receptor, type 1' FT
FT 1403..2028.
FEATURES
3'UTR Location/Qualifiers
source 1..2028
/organism="Homo sapiens"
/db_type="genomic RNA"
/mol_xref="taxon:9606"
ORIGIN
Alignment Scores: 5.58e-65 Length: 2028
Pred. No.: 1042.50 Matches: 197
Score: 87.2% Conservative: 41
Percent Similarity: 72.2% Mismatches: 34
Best Local Similarity: 50.6% Indels: 1
Query Match: 6 Gaps: 1
DB:
US-10-618-570-2 (1-400) x E03267 (1-2028)
Qy 1 MetAlaGlnTrpAspAspPheProAspGlnGlnGluAspThrAspSerCysThrGluSer 20
Db 47 ATGAGCGTGGGATCACTTTCAATCAACAGAGGACACTGATAGCTCTCGAATCT 106
Qy 21 ValLysPheAspAlaArgSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
Db 107 GTGAATTTGATGCTCGCTCAATGACAGAGCTTTGCTTCTCCGAATCTTAAACACAGCCT 166
Qy 41 ThrLeuGlnGluArgMetLysSerTyrLysThrAlaLeuIleThrLeuTyrLeuIleVal 60
Db 167 TCCCTTCAAGAGAACTGAAGTCTCTCAAAAGCTGCACCTGATTGCCCTTTTACCTCTCGTG 226
Qy 61 PheValValLeuValProIleIleGlyIleValAlaAlaGlnLeuLeuValTrpGluThr 80
Db 227 TTTGCAGTTCTCATCCCTCTCATTTGGATAGTGGCAGCTCACTCTCTCAAGTGGGAACG 286
Qy 81 LysAsnCysThrValGlySerValAsnAla---AspIleSerProSerProGluGlyLys 99
Db 287 AAGAATTTGCTCAGTTAGTTCAACTAATGCAATGATATAACTCAAGTCTCACGGGAAA 346
Qy 100 GlyAsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsn 119
Db 347 GGAATGACAGCGAGAGGAATGAGATTTCAAGAACTCTTTATGGAACACATGAGCAAC 406
Qy 120 MetGluSerArgIleGlnTyrLeuSerAspAsnGlnAlaAsnLeuLeuAspAlaLysAsn 139
Db 407 ATGGAAGAAGATCCAGCATATTTAGACATGGNAGCCAACTCATGGACACAGAGCAT 466
Qy 140 PheGlnAsnPheSerIleThrThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsn 159
Db 467 TTCCAAAATTTTCAGCATGACAACTGATCAAGATTTAATGACATTTCTTCTGCAGCTAAGT 526
Qy 160 SerLeuLeuSerSerIleGlnGluHisGluAsnIleIleGlyAspIleSerLysSerLeu 179
Db 527 ACCTTGTGTTTCTCAGTCCAGGAGCATGGGAATGCAATAGATGAAATCTCCAAGTCTCTTA 586
Qy 180 ValGlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArg 199
Db 587 ATAAGTTTGAATACCATTTGCTTGTATTGACGCTCAACATAGAAAATCTGAATGGCAAA 646
Qy 200 ValGlnGluAsnAlaPheLysGlnGlnGluMetArgLysLeuGluGluArgIleTyr 219
Db 647 ATCCAAGAGATACCTTCAACACACAGAGGAAATCAGTAATTTAGAGGCGGTGTTCAC 706
Qy 220 AsnAlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyrLeuGluGlnGluIle 239
Db 707 AATGATATCAGCAGAAATTTATGGCTATGAAAGAGAAACAAAGTGCAATTTGGAAACAGGAATA 766

```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 27, 2006, 09:14:37 ; Search time 5148 Seconds
(without alignments)
3635.360 Million cell updates/sec

Title: US-10-618-570-2

Perfect score: 2059

Sequence: 1 MAQWDDPPQDQEDTDSCTES.....DWKATRVGINIFRLRTOKE 400

Scoring table:

| | |
|---------------------------|---------------------------|
| BLOSUM62 | Xgapop 10.0 , Xgapext 0.5 |
| Ygapop 10.0 , Ygapext 0.5 | |
| Fgapop 6.0 , Fgapext 7.0 | |
| Delop 6.0 , Delext 7.0 | |

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/abs/ABSSWEB.spool/US10618570/runat.27022006.063856.13092/app.query.fasta_1
-DB=EST -QFMT=fastap -SUFFIX=rat -MINMATCH=0.1 -LOPCL=0 -LOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORES=500 -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZES=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs807
-USER=US10618570 @CGN 1 1 5315 @runat.27022006.063856.13092 -NCPU=6 -ICPU=3
-NO.MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*

| Query | Match | Length | ID | Description |
|----------------|-------|--------|----|-------------|
| 1: gb_est1:* | 62.0 | 793 | 8 | DN548862 |
| 2: gb_est2:* | 55.4 | 949 | 6 | CF409555 |
| 3: gb_est3:* | 50.6 | 1033 | 10 | AY413208 |
| 4: gb_est4:* | 49.5 | 652 | 6 | CB453082 |
| 5: gb_est5:* | 48.7 | 1027 | 10 | AY413209 |
| 6: gb_est6:* | 45.9 | 978 | 7 | CN643776 |
| 7: gb_est7:* | 43.0 | 843 | 5 | EX496563 |
| 8: gb_est8:* | | | | |
| 9: gb_est9:* | | | | |
| 10: gb_est10:* | | | | |
| 11: gb_est11:* | | | | |

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Length | ID | Description |
|------------|--------|-------|--------|----|-------------|
| 1 | 1276 | 62.0 | 793 | 8 | DN548862 |
| 2 | 1140.5 | 55.4 | 949 | 6 | CF409555 |
| 3 | 1042.5 | 50.6 | 1033 | 10 | AY413208 |
| 4 | 1020 | 49.5 | 652 | 6 | CB453082 |
| 5 | 1002.5 | 48.7 | 1027 | 10 | AY413209 |
| 6 | 944.5 | 45.9 | 978 | 7 | CN643776 |
| 7 | 885 | 43.0 | 843 | 5 | EX496563 |

| 8 | 830 | 40.3 | 3635 | 4 | AK028480 |
|----|-------|------|------|----|----------|
| 9 | 823 | 40.0 | 1030 | 10 | AY413210 |
| 10 | 801 | 38.9 | 1509 | 4 | AK089178 |
| 11 | 791.5 | 38.4 | 677 | 5 | BW956649 |
| 12 | 768.5 | 37.3 | 768 | 8 | DR004903 |
| 13 | 736.5 | 35.8 | 626 | 6 | CB469554 |
| 14 | 727.5 | 35.3 | 651 | 8 | DR004296 |
| 15 | 719.5 | 34.9 | 1093 | 7 | CN643736 |
| 16 | 717 | 34.8 | 816 | 6 | CB997441 |
| 17 | 714.5 | 34.7 | 665 | 6 | CB472122 |
| 18 | 698.5 | 33.4 | 797 | 6 | CB990950 |
| 19 | 678 | 32.9 | 559 | 6 | CF251702 |
| 20 | 678 | 32.9 | 561 | 6 | CD725634 |
| 21 | 678 | 32.9 | 575 | 5 | BU447766 |
| 22 | 678 | 32.9 | 576 | 5 | BU408647 |
| 23 | 678 | 32.9 | 582 | 5 | BX269637 |
| 24 | 678 | 32.9 | 586 | 6 | CF251883 |
| 25 | 678 | 32.9 | 588 | 6 | CF252572 |
| 26 | 678 | 32.9 | 594 | 6 | CF252815 |
| 27 | 674 | 32.7 | 581 | 6 | CF251833 |
| 28 | 673 | 32.7 | 576 | 5 | BU447569 |
| 29 | 670 | 32.5 | 570 | 5 | BU409120 |
| 30 | 670 | 32.5 | 578 | 5 | BU426481 |
| 31 | 670 | 32.5 | 612 | 5 | BU422332 |
| 32 | 669.5 | 32.5 | 585 | 5 | BW958294 |
| 33 | 665 | 32.3 | 594 | 6 | CF252175 |
| 34 | 664 | 32.2 | 538 | 3 | BM489973 |
| 35 | 657 | 31.9 | 473 | 3 | BQ038085 |
| 36 | 657 | 31.9 | 693 | 1 | AW744595 |
| 37 | 654 | 31.8 | 589 | 6 | CF251795 |
| 38 | 649 | 31.5 | 453 | 7 | CK608106 |
| 39 | 649 | 31.5 | 458 | 7 | CK614089 |
| 40 | 647 | 31.4 | 456 | 7 | CK614219 |
| 41 | 647 | 31.4 | 721 | 6 | CF745486 |
| 42 | 645 | 31.3 | 460 | 7 | CK615416 |
| 43 | 641 | 31.1 | 460 | 7 | CK608142 |
| 44 | 641 | 31.1 | 465 | 7 | CK614950 |
| 45 | 637 | 30.9 | 593 | 5 | BU418594 |

ALIGNMENTS

RESULT 1

DN548862

LOCUS

DEFINITION

ACCSSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

DN548862 1407127 MARC 7BOV Bos taurus CDNA 5', mRNA sequence. EST 11-MAR-2005

DN548862.1 GI:61013550

EST.

Bos taurus (cow)

Bos taurus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;

Pecora; Bovidae; Bovinae; Bos.

Smith, T.P.L., Roberts, A.J., Echtenkamp, S.E., Chitko-Mckown, C.G.,

Wray, J.E. and Keeler, J.W.

A second set of bovine ESTs from pooled-tissue normalized libraries

Unpublished (2003)

Contact: Smith TPL

USDA, ARS, US Meat Animal Research Center

PO Box 166, Clay Center, NE 68933-0166, USA

Tel: 402 762 4366

Fax: 402 762 4390

Email: smith@email.marc.usda.gov

Single pass sequencing. Bases called with phred v0.020425.c and

trimmed with the aid of the trim_alt option. Vector identified with

cross match v0.990329.

Plate: QOP8005 row: C column: 19

Seq primer: GTAATACGACTACTATAGG.

Location/Qualifiers

1. .793

/organism="Bos taurus"

FEATURES

source

```

/mo1_type="mRNA"
/db_xref="taxon:9913"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 7B0V"
/notes="Vector: pcDNA3.1; Site 1: EcoRI; Site 2: NotI;
Library made with RNA pooled from multiple tissues
including ovary, hindbrain, uterus, and day-30 whole
embryos."

ORIGIN
Alignment Scores:
Pred. No.: 1,24e-138 Length: 793
Score: 1276.00 Matches: 252
Percent Similarity: 100.0% Conservatve: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 62.0% Indels: 0
DB: 8 Gaps: 0

US-10-618-570-2 (1-400) x DN548862 (1-793)

Qy 1 MetAlaGlnTrpAspPheProAspGlnGlnGluAspThrAspSerCysThrGluSer 20
Db ATGGCACAGTGGGATGACTTTCCTGATCAGCAAGAGGACACTGACAGCTGTACAGAGTCT 95
Qy 21 VallysPheAspAlaArgSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
Db GTGAAGTTCGATCGCTCGCTCAGTGACAGCTTTGGCTTCTCCCTCCCTAAATAATGGGCCA 155
Qy 41 ThrLeuGlnGluArgMetLysSerTyrLysThrAlaLeuLeuThrLeuTyrLeuVal 60
Db ACTCTTCAAGAGAGATGAGTCTTATTAATACTGCATGCTGATCCCTTATCTCATTTGTG 215
Qy 61 PheValValLeuValProileGlyLeValAlaAlaGlnLeuLeuLysTyrTrpGluThr 80
Db TTTGTAGTTCCTGTCGCCCATCTTGGCATAGTGCGCAGCTCAGCTCCTGAAATGGGAAACG 275
Qy 81 LysAsnCysThrValGlySerValAsnAlaAspLysSerProSerProGluGlyLysGly 100
Db AAGATTTCGACGGTTGGCTCAGTTAATGCAGATATATCTCAAGTCCGGAAGGCAAGGA 335
Qy 101 AsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsnMet 120
Db AATGCGATGGAATGAATGAGATTCGAGAGCTGTGATGGAACGATGAGCAACATG 395
Qy 121 GluSerArgIleGlnTyrLeuSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsnPhe 140
Db GAAAGCAGATCCAGTATCTTTCAGATTAATGAAGCCAATCTCCTAGATGCTAAGAAATTC 455
Qy 141 GlnAsnPheSerIleThrThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsnSer 160
Db CAATAATTCAGCATAAACAACTGATCAAAAGATTTAATGATGTCTTTTCCAGCTAAATTC 515
Qy 161 LeuLeuSerSerIleGlnGluHisGluAsnIleLeGlyAspIleSerLysSerLeuVal 180
Db TTACTTTCTCCATCCAGAACATGAGAAATATCATAGGGGATATCTCCAACTCATAGTA 575
Qy 181 GlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArgVal 200
Db GGTCTGAACACCCACAGTACTTGGATTTGCGATTCAGTATTTGAACACACTGAATGGCAGAGTC 635
Qy 201 GlnGluAsnAlaPheLysGlnGlnGluMetArgLysLeuGluGluArgIleTyrAsn 220
Db CAAGAGAAATGCAATTTAAACAAACAGAGGAGATGCGTAAATTTAGAGGCGGTATATACAAT 695
Qy 221 AlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyrLeuGluGlnGluLys 240
Db GCATCAGCAGAAATTAAGTCTCTAGATGAAACAAAGTATATTTGGAAACGAAATATAA 755
Qy 241 GlyGluMetLysLeuLeuAsnAsnIleThrAsnAsp 252
Db GGGGAAATGAATGTTGTAATAATATATCACTAATGAT 791

```

RESULT 2
CF409555 949 bp mRNA linear EST 02-SEP-2003
LOCUS CH3#061_D06MF Canine heart normalized cDNA Library in pBluescript
DEFINITION Canis familiaris cDNA clone CH3#061_D06 5', mRNA sequence.
ACCESSION CF409555
VERSION CF409555.1 GI:34410801
KEYWORDS EST.
SOURCE Canis familiaris (dog)
ORGANISM Canis familiaris
REFERENCE 1 (bases 1 to 949)
AUTHORS Yi,Y., Desai,R., Olarte,M., Henthorn,P. and George A.L.
TITLE Expressed sequence tags from Canine heart
JOURNAL Unpublished (2003)
COMMENT Other ESTs: CH3#061_D06MR
Contact: George AL_D06MR
Division of Genetic Medicine
Vanderbilt University
529 Light Hall, 2215 Garland Avenue, Nashville, TN 37232-0275, USA
Tel: 615 936 2660
Fax: 615 936 2661
Email: al.george@vanderbilt.edu
Insert Length: 2287 Std Error: 0.00
Seq primer: MF: GTTTCCTCCAGTCACGACGTG
High quality sequence start: 169
High quality sequence stop: 784.
Location/Qualifiers
1..949
/organism="Canis familiaris"
/mol_type="mRNA"
/db_xref="taxon:9615"
/clone="CH3#061_D06"
/tissue_type="heart"
/cell_type="heart"
/dev_stage="mixed developmental stages (adult, 30 day - 40 day fetal)"
/clone_lib="Canine heart normalized cDNA Library in pBluescript"
/notes="Organ: heart; Vector: pBluescript; Site 1: 5' of vector NotI; Site 2: 3' of vector EcoRI; Tissue source: dog heart (adult, 30 day - 40 day fetal), right and left atria and ventricle. Dog breed - mixed (beagle, German shepherd, pointer, Irish setter). Library construction: oligo-dT primed"

ORIGIN
Alignment Scores:
Pred. No.: 1,32e-122 Length: 949
Score: 1140.50 Matches: 235
Percent Similarity: 96.0% Conservatve: 3
Best Local Similarity: 94.8% Mismatches: 6
Query Match: 55.4% Indels: 4
DB: 6 Gaps: 3

US-10-618-570-2 (1-400) x CF409555 (1-949)
Qy 1 MetAlaGlnTrpAspPheProAspGlnGlnGluAspThrAspSerCysThrGluSer 20
Db ATGGCACAGTGGGATGACTTTCCTGATCAGCAAGAGGACACTGACAGCTGTACAGAGTCT 264
Qy 21 VallysPheAspAlaArgSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
Db GTGAAGTTCGATCGCTCGCTCAGTGACAGCTTTGGCTTCTCCCTCCCTAAATAATGGGCCA 324
Qy 41 ThrLeuGlnGluArgMetLysSerTyrLysThrAlaLeuLeuThrLeuTyrLeuVal 60
Db ACTCTTCAAGAGAGATGAGTCTTATTAATACTGCATGCTGATCCCTTATCTCATTTGTG 384
Qy 61 PheValValLeuValProIleGlyLeValAlaAlaGlnLeuLeuLysTyrTrpGluThr 80

```

Db      385 TTTGTAGTTCTCGTCCCATCTGGCATAGTGGCAGCTCAGCTCCTGAAATGGGAACG 444
Qy      81 LysAnCysThrValGlySerValAsnAlaAspIleSerProSerProGluGlyIleGly 100
Db      445 AAGAATTGACGGTGGCTGAGTTAATGACATATATCTCCAACTGCGGAAGGCAAGGA 504
Qy      101 AsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsnMet 120
Db      505 AATGGCAGTGAAGATGAATGAGATTTTCGAGAAGCTGTGATGGAAACGATGAGCAACATG 564
Qy      121 GluSerArgIleGlnTyrLeuSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsnPhe 140
Db      565 GAAAGCAGATCCAGTATCTTTCCAGATAATGAAGCCAACTCTCTAGATGCTAAGATTTTC 624
Qy      141 GlnAsnPheSerIleThrThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsnSer 160
Db      625 CAAAATTTCCAGATAACAACTGATCAAGATTTAATGATGTTCTTTCCAGCTAAATTC 684
Qy      161 LeuLeuSerSerIleGlnGluHieGluAsnIleIleGlyAspIleSerLysSerLeuVal 180
Db      685 TTACTTTCTCCATCCAGAACATGAGAAATATCATAGGGATATCTCCAAAGTCATTAGTA 744
Qy      181 GlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArgVal 200
Db      745 GGTCTGACACACACAGTACTTGTATTTGCGATTCAGTATTGAAACACTGAATGGCAGATC 804
Qy      201 GlnGluAsnAlaPheLysGlnGlnGluMetArgLysLeuGluGluArgIleTyrAsn 220
Db      805 CAAGAGAATGCATTTAAACCAACAGAGAGATGCGTAAATTTAGAGGCGGTATATACAAT 864
Qy      221 AlaSerAlaGluLleLysSerLeuAspGluLys---GlnValTyrIleu---GluGlnG 238
Db      865 GCATCAGCAGAAAATTAAGTCTCTAGATGGAACCAACCAAGTATATTTTGGGAACAGGA 924
Qy      238 uileLys---GlyGluMetLys 244
Db      925 AATTAAAGGGGAAAATGGAA 946

RESULT 3
AY413208
LOCUS      1033 bp      DNA      linear      GSS 17-DEC-2003
DEFINITION Homo sapiens MSR1 gene, VIRTUAL TRANSCRIPT, partial sequence,
GENOMIC SURVEY SEQUENCE.
ACCESSION AY413208
VERSION   AY413208.1 GI:39769173
KEYWORDS GSS.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
          Homnidae; Homo.
          1 (bases 1 to 1033)
          Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
          Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,P., Murphy,B.,
          Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
          Adams,M.D. and Cargill,M.
          Inferring nonneutral evolution from human-chimp-mouse orthologous
          gene trios
          Science 302 (5652), 1960-1963 (2003)
          14671302
          2 (bases 1 to 1033)
          Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
          Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,P., Murphy,B.,
          Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
          Adams,M.D. and Cargill,M.
          Direct Submission
          Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
          Rockville, MD 20850, USA
          This sequence was made by sequencing genomic exons and ordering
          them based on alignment.
          Location/Qualifiers
          1..1033
          /organism="Homo sapiens"

```

```

/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>1033
/gene="MSR1"
/locus_tag="HCM4814"

ORIGIN

Alignment Scores:
Pred. No.:      4.82e-111      Length:      1033
Score:          1042.50      Matches:      197
Percent Similarity: 87.2%      Conservative: 41
Best Local Similarity: 72.2%      Mismatches: 34
Query Match:    50.6%      Indels:      1
Gaps:           10           Gaps:         1

US-10-618-570-2 (1-400) x AY413208 (1-1033)

Qy      1 MetAlaGlnTrpAspAspPheProAspGlnGlnGluAspThrAspSerCysThrGluSer 20
Db      1 ATGGAGCAGTGGGATCACTTTTCACATCAACAGAGGACACTGATAGCTGCTCCGAATCT 60
Qy      21 ValLysPheAspAlaArgSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
Db      61 GTGAATTTGATGCTCGCTCAATGACAGCTTTGCTTCTCCGATCTTAAACACAGCCCT 120
Qy      41 ThrLeuGlnGluArgMetLysSerTyrLysThrAlaLeuIleThrLeuTyrLeuIleVal 60
Db      121 TCCCTTCAAGAGAACTGAAGTCTCTCAATGACAGCTTTGCTTCTCCGATCTTAAACACAGCCCT 180
Qy      61 PheValValLeuValProIleIleGlyIleValAlaAlaGlnLeuLeuLysTyrGluThr 80
Db      181 TTTGCACTTCTCATCCCTCTCATTTGGAATAGTGGCAGCTCAACTCTGGAAGTGGGAACG 240
Qy      81 LysAsnCysThrValGlySerValAsnAla---AspIleSerProSerProGluGlyLys 99
Db      241 AAGBATTGCTCAGTTAGTTCACTAATGCAATGATATATACTCAAGTCTCACGGGAAA 300
Qy      100 GlyAsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsn 119
Db      301 GGAATGACAGCGAAGAGAAATGAGATTTCAAGAGCTTTTATGGAACACATGAGCAAC 360
Qy      120 MetGluSerArgIleGlnTyrLeuSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsn 139
Db      361 ATGGAGAAGAGAAATCCAGCATATTTTAGACATGATGAGCCCAACCTCATGGACACAGAGAT 420
Qy      140 PheGlnAsnPheSerIleThrThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsn 159
Db      421 TTCCAAATTTTCCAGTACGACACTGATCAAGATTTAATGACATCTTCTGCAGCTAAGT 480
Qy      160 SerLeuLeuSerSerIleGlnGluHisGluAsnIleIleGlyAspIleSerLysSerIleu 179
Db      481 ACCTTGTTTCTCAGTCCAGGGACATGGGAATGCAATAGATGAATCTCCAGTCTCTTA 540
Qy      180 ValGlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArg 199
Db      541 ATAAGTTTGAATACCACTGCTTGAATTTGAGTTCAGCTCAACATAGAAAATCTGAATGGCAAA 600
Qy      200 ValGlnGluAsnAlaPheLysGlnGlnGluMetArgLysLeuGluGluArgIleTyr 219
Db      601 ATCCAGAGAAATACCTTCAAAACACAGAGAAATCATGTAATTTAGAGGACGGTGTAC 660
Qy      220 AsnAlaSerAlaGluIleLysSerLeuAspGlnLysGlnValTyrLeuGluGlnGluIle 239
Db      661 AATGATATCAGCAGAAATATTATGGCTATGAAAGAGAAACAGTGCATTTTGGAAACAGGAATA 720
Qy      240 LysGlyGluMetLysLeuLeuAsnAsnIleThrAsnAspLeuArgLeuLysAspTyrGlu 259
Db      721 AAAGGAGAGTGAAGTACTGAAATACATCACTAATGATCTCAGACTGAAGATTTGGGAA 780
Qy      260 HisSerGlnThrLeuLysAsnIleThrLeuLeuGlnGly 272
Db      781 CATTTCTCAGACCTTGAAGAAATATACCTTTAATTAACAGGT 819

```

```

RESULT 4
CB453082          652 bp      mRNA      linear      EST 26-MAR-2003
LOCUS             708007 MARC 6BOV Bos taurus cDNA 5', mRNA sequence.
DEFINITION        CB453082
ACCESSION         CB453082
VERSION           CB453082.1 GI:29259464
KEYWORDS          EST.
SOURCE            Bos taurus (cow)
ORGANISM          Bos taurus
                  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                  Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
                  Pecora; Bovidae; Bovinae; Bos.
REFERENCE         1 (bases 1 to 652)
AUTHORS           Smith,T.P.L., Roberts,A.J., Ehternkamp,S.E., Chitko-McKown,C.G.,
                  Way,J.E. and Keele,J.W.
TITLE             A second set of bovine ESTs from pooled-tissue normalized libraries
JOURNAL           Unpublished (2003)
COMMENT           USDA, ARS, US Meat Animal Research Center
                  PO Box 166, Clay Center, NE 68933-0166, USA
                  Tel: 402 762 4366
                  Fax: 402 762 4390
                  Email: smith@mail.marc.usda.gov
                  Single pass sequencing. Bases called with phred v0.020425.c and
                  trimmed with the aid of the trim_alt option. Vector identified with
                  cross_match v0.990329.
                  Plate: FQY8065 row: M column: 21
                  Seq primer: GTAATGACACTCATATAGGG.
FEATURES          source
                  1..652
                  /organism="Bos taurus"
                  /mol_type="mRNA"
                  /db_xref="taxon:9913"
                  /tissue_type="pooled"
                  /lab_host="DH10B"
                  /clone_lib="MARC 6BOV"
                  /note="Vector: pCDNA3.1; Site 1: EcoRI; Site 2: NotI;
                  Library made with RNA pooled from multiple tissues
                  including liver, lung, hypothalamus, pituitary, and
                  placenta/endometrium."
ORIGIN
Alignment Scores:
Pred. NO.:        1.05e-108      Length:      652
Score:            1020.00        Matches:    207
Percent Similarity: 81.8%      Conservative: 0
Best Local Similarity: 81.8%    Mismatches: 0
Query Match:      49.5%        Indels:     46
DB:               6            Gaps:        1

US-10-618-570-2 (1-400) x CB453082 (1-652)

Qy      1 MetAlaGlnTrpAspPheProAspGlnGlnGluAspThrAspSerCysThrGluSer 20
Db      32 ATGGCACAGTGGGATGACCTTTCCTGATCAGCAAGAGGACACTGACAGCTGTACAGAGTCT 91
Qy      21 ValLysPheAspAlaArgSerValThrAlaLeuLeuProProHisProLysAsnGlyPro 40
Db      92 GTGAAGTTTCGATGCTCGCTCAGTGACAGCTTTGCTTCTCCCATCTCTAAAAATGGCCCA 151
Qy      41 ThrLeuGlnGluArgMetLysSerTyrLysThrAlaLeuLeuThrLeuTyrLeuLeuVal 60
Db      152 ACTCTTCAAGAGAGATGAGTCTTATTAATACTGCAGCTGATCACCCCTTATCTCATTTGTG 211
Qy      61 PheValValLeuValProIleGlyIleValAlaAlaGlnLeuLeuLysTrpGluThr 80
Db      212 TTTGTAGTTCTCGTCCCATCATCTTGGCATAGTGGCAGCTCAGCTCCTGAAATGGAAACG 271
Qy      81 LysAsnCysThrValGlySerValAsnAlaAspIleSerProSerProGluGlyLysGly 100
Db      272 AAGAATTGCACGGTTGGCTCAGTTAATGCAGATATATCTCCAGTCCGGAAGGCAAGGA 331
Qy      101 AsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsnMet 120

```

```

Db      332 AATGGCAGTGAACATGAATGAGATTCGAGAAGCTGTGATGGAACGATGAGCAACATG 391
Qy      121 GluSerArgIleGlnTyrLeuSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsnPhe 140
Db      391 -----
Qy      141 GlnAsnPheSerIleThrThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsnSer 160
Db      391 -----
Qy      161 LeuLeuSerSerIleGlnGluHisGluAsnIleIleGlyAspIleSerLysSerLeuVal 180
Db      392 -----GAACATGAGATATCATAGGGGATATCTCCAAGTCATTAGTA 433
Qy      181 GlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGlnThrLeuAsnGlyArgVal 200
Db      434 GGTCTGAACACCAACAGTACTTGTGATTTGCAGTTTCAGTATTGAAACACTGAATGCGAGATC 493
Qy      201 GlnGluAsnAlaPheLysGlnGlnGluMetArgLysLeuGluGluArgIleTyrAsn 220
Db      494 CAAGAGATGCATTTAAACAACAAGAGAGATCGTAAATTAGAGGAGCGTATATACAT 553
Qy      221 AlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyrLeuGluGlnGluIleLys 240
Db      554 GCATCAGCAGAGAAATTAAGTCTCTAGATGAAAAACAAGTATATTTGGAACAGGAATAAAA 613
Qy      241 GlyGluMetLysLeuLeuAsnAsnIleThrAsnAspLeu 253
Db      614 GGGGAAATGAACCTGTTGAATAATATCACTAATGATCTG 652

RESULT 5
AY413209          1027 bp      DNA      linear      GSS 17-DEC-2003
LOCUS             Pan troglodytes MSRI gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION        genomic survey sequence.
ACCESSION         AY413209
VERSION           AY413209.1 GI:39769174
KEYWORDS          GSS.
SOURCE            Pan troglodytes (chimpanzee)
ORGANISM          Pan troglodytes
                  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                  Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
                  Homiidae; Pan.
REFERENCE         1 (bases 1 to 1027)
AUTHORS           Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
                  Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,P., Murphy,B.,
                  Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
                  Adams,M.D. and Cargill,M.
TITLE             Inferring nonneutral evolution from human-chimp-mouse orthologous
                  gene trios
JOURNAL           Science 302 (5652), 1960-1963 (2003)
PUBMED           14671302
REFERENCE         2 (bases 1 to 1027)
AUTHORS           Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
                  Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,P., Murphy,B.,
                  Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
                  Adams,M.D. and Cargill,M.
TITLE             Direct Submission
JOURNAL           Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
                  Rockville, MD 20850, USA
COMMENT           This sequence was made by sequencing genomic exons and ordering
                  them based on alignment.
FEATURES          Location/Qualifiers
                  1..1027
                  /organism="Pan troglodytes"
                  /mol_type="genomic DNA"
                  /db_xref="taxon:9598"
                  <1..>1027
                  /gene="MSRI"
                  /locus_tag="HCM4814"
ORIGIN

```


[illegible]


```

Db      835 ACATCT 841
|||||
RESULT 8
AK028480
LOCUS   3636 bp mRNA linear HTC 03-APR-2004
DEFINITION Mus musculus 0 day neonate skin cDNA, RIKEN full-length enriched library, clone:4631424C05 product:macrophage scavenger receptor 1, full insert sequence.
ACCESSION AK028480
VERSION   GI:26324437
KEYWORDS  HTC; CAP trapper.
SOURCE    Mus musculus (house mouse)
ORGANISM  Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
10349636

2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
11042159

3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Kitsu, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kaishiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
11076861

4
The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)

5
The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 3636)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirao, T., Hirozane, T., Hori, F., Imoto, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saitoh, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, P., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gs.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

```

COMMENT

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: <http://genome.gsc.riken.jp/>

URL: <http://fantom.gsc.riken.jp/>

FEATURES

Location/Qualifiers

source

1..3636
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="FANTOM DB:4631424C05"
 /db_xref="taxon:10090"
 /clone="4631424C05"
 /tissue_type="skin"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="0 day neonate"

CDS

53..1117
 /note="unnamed protein product; macrophage scavenger receptor 1 (MGD) [MGI:98257, GB|NM_031195, evidence: BLASTN, 99% match=1461]"
 putative"
 /codon_start=1
 /protein_id="BAC25973.1"
 /db_xref="GI:26324438"
 /translation="MTKMTENQRLCPHEQEDADCSSESVKFADRSMTASLPHSTKNG
 PSLOELKSPKAAIALYLVFVLPVGVITAGLLNWMKNCGLVCSINTSDTSQGP
 MEKTSKVENRFTLIMEHKMDSERIESINSKADLIDTERFQNFQSWATDORLDIL
 LQNSLSSVQEHGNSDAISQSLNMLLDVLTETLNLRVRESATKQEDISK
 LEERYKVSRAEVQKEQAEVQEVRLNNITNDLRKQWHSQTLKNTFFIQ
 GPPQGEKGRGLTGTGTPGPGICIRGIPGVGRDQIGFPGGRGNPCAGPKGRSG
 SPGPKQKGKSGSVGSSRSV"

ORIGIN

Alignment Scores:
 Pred. No.: 2,93e-85 Length: 3636
 Score: 830.00 Matches: 162
 Percent Similarity: 78.6% Conservative: 47
 Best Local Similarity: 60.9% Mismatches: 55
 Query Match: 40.3% Indels: 2
 DB: 4 Gaps: 2

US-10-618-570-2 (1-400) x AK028480 (1-3636)

Qy 8 ProaspGlnGlnGluAaspThrAspSerCysThrGluSerVallyPheAspAlaArgSer 27
 Db 89 CCTCATGAACAAAGAGGATGCTGACTGCAGTTCAGAAATTTGACGACGTTCA 148
 Qy 28 ValThrAlaLeuLeuProHisProLysAsnGlyProThrLeuGlnGluArgMetLys 47
 Db 149 ATGACTGCATCCCTTCCTCAGCAGCACTAAATGGCCCTTCCTTCAGAGAGTTGAAG 208
 Qy 48 SerTyrIysThrAlaLeuLeuThrLeuTyrLeuLeuValPheValLeuValProIle 67
 Db 209 TCCTTCAAGGCTGCCCTCATTTGCTCTACCTCTCTGTTTGGCAGTACTAATACCTGTT 268
 Qy 68 IledlylleValAlaAlaGlnLeuLeuLysTrpGluThrLysAsnCysThrValGlySer 87
 Db 269 GTTGGAAATAGTAACAGCTCAGCTTTTGAATTTGGGAAATGAAGAACTGTTAGTTTGTCA 328
 Qy 88 ValAsn---AlaAapIleSerProSerProGluGlyLysGlyAsnGlySerGluAaspGlu 106
 Db 329 CTTAAACAAGTGACACATCTCAAGGTCCTATGGAAAGAAAGAAATACCAAGTAAGTGGAA 388
 Qy 107 MetArgPheArgGluAlaValMetGluArgMetSerAsnMetGluSerArgIleGlnTyr 126
 Db 389 ATGAGATT---ACAATTATCATGACACATGAAGGACATGAGGAGGAGAGATCGAAGC 445
 Qy 127 LeuSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsnPheGlnAsnPhSerIleThr 146
 Db 446 ATTTCAAACTCAAAAGCGGACCTTTATACACACGGAACGCTTCCAGAAATTTCCAGCATGCA 505

Qy 147 ThrAspGlnArgPheAsnAspValLeuPheGlnLeuAenSerLeuLeuSerIleGln 166
 Db 506 ACTGACCAAGACTTAAATGATGATCTCTGAGTTAAATTCCTTGATTCGTGAGTCAG 565

Qy 167 GluHisGluAenIleGlyAspIleSerLeuValGlyLeuAenThrVal 186
 Db 566 GAACATGGGAATTCACCTGGATGCATCTCCAAAGTCCTTCGAGAGTCGTAATATGACACTG 625

Qy 187 LeuAspLeuGlnPheSerIleGlnThrLeuAenGlyArgValGlnGlnAenAlaPheLys 206
 Db 626 CTTGATGTTCAACTCATACAGAACTGAATGTGACAGTCCGTGTAATCTACAGCAAG 685

Qy 207 GlnGlnGluGluMetArgLysLeuGluGluArgLyleTyrAenAlaSerAlaGluLileLys 226
 Db 686 CAACAGSAGGACATCAGTAAATGGAGAACTGTGTACAAAGTATACGAGAGTCCAG 745

Qy 227 SerLeuAspGluLysGlnValTyrLeuGluGlnGluLileLysGlyGluMetLysLeuLeu 246
 Db 746 TCTGTGAAGAGAACACAGCGCAGCTGGAAACAGGAGTAAACAGGAGTGAAGTATTG 805

Qy 247 AenAenIleThrAsnAspLeuArgLysLeuLysAspTrpGluHisSerGlnThrLeuLysAen 266
 Db 806 AACACATCACCACAGCTCAGACTGAAGAGTGGGAACACTCAGACACTGAAAAAC 865

Qy 267 IleThrLeuLeuGlnGly 272
 Db 866 ATCACCCTTCATTCAGGG 883

RESULT 9
 AY413210 1030 bp DNA linear GSS 17-DEC-2003
 LOCUS Mus musculus MSRI gene, VIRUAL TRANSCRIPT, partial sequence,
 DEFINITION genomic survey sequence.
 ACCESSION AY413210
 VERSION AY413210.1 GI:39769175
 KEYWORDS GSS
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridae; Muridae; Mus.
 REFERENCE 1 (bases 1 to 1030)
 AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
 Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
 Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
 Adams,M.D. and Cargill,M.
 TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
 gene trios
 JOURNAL Science 302 (5652), 1960-1963 (2003)
 PUBMED 14671302
 REFERENCE 2 (bases 1 to 1030)
 AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
 Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
 Perriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
 Adams,M.D. and Cargill,M.
 TITLE Direct Submission
 JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
 Rockville, MD 20850, USA
 COMMENT This sequence was made by sequencing genomic exons and ordering
 them based on alignment.
 FEATURES
 Location/Qualifiers
 1..1030
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 <1..>1030
 /gene="MSRI"
 /locus_tag="HCM4814"

ORIGIN
 Alignment Scores: 2.84e-85 Length: 1030
 Pred. No.:

Score: 823.00 Matches: 160
 Percent Similarity: 78.6% Conservative: 49
 Best Local Similarity: 60.2% Mismatches: 55
 Query Match: 40.0% Indels: 2
 DB: 10 Gaps: 2

US-10-618-570-2 (1-400) x AY413210 (1-1030)

Qy 8 ProAspGlnGlnGluAenThrAspSerCysThrGluSerValLysPheAsnAlaArgSer 27
 Db 22 CCTCATGACGAGAGGATGCTGACTGCAGTTCAAGTCCGTAATTTTGACGCGACGTTCA 81

Qy 28 ValThrAlaLeuProHisProLysAenGlyProThrLeuGlnGluArgMetLys 47
 Db 82 ATGACAGATCCCTTCTCTCACAGCACTAAATATGCCCCCTCCGTTCAGGAGAGTTGAAG 141

Qy 48 SerTyrLysThrAlaLeuIleThrLeuTyrLeuIleValPheValValLeuValProIle 67
 Db 142 TCCTTCAAGGCTGCCCTCATTTGCTCTACCTCTGTGTTTGCAGTACTAATACCTGTT 201

Qy 68 IleGlyIleValAlaAlaGlnLeuLeuLysTrpGluThrLysAenCysThrValGlySer 87
 Db 202 GTTCGAATAGTCAGCTCAGCTTTTGAATTTGGGAAATGAAGAACTGCTTAGTTGTTCA 261

Qy 88 ValAen--AlaAspIleSerProSerProGluGlyLysGlyAenGlySerGluAspGlu 106
 Db 262 CTTAACACAGTGACACATCTCAAGTCTCTATGAAAGAAAGAAATACCACTAAAGTGAA 321

Qy 107 MetArgPheArgGluAlaValMetGluArgMetSerAenMetGluSerArgIleGlnTyr 126
 Db 322 ATGAGATTT--ACAATTTATCATGGAACACATGAAGGACATGGAGGAGAGAATCGAAAGC 378

Qy 127 LeuSerAspAenGluAlaAenLeuLeuAspAlaLysAenPheGlnAenPheSerIleThr 146
 Db 379 ATTTCAAACTCAAAGCCGACCTTATAGACACGGAACGCTTCCAGAAATTCAGCATGCGCA 438

Qy 147 ThrAspGlnArgPheAenAspValLeuPheGlnLeuAenSerLeuLeuSerSerIleGln 166
 Db 439 ACTGACCAAGACTTAATGATATTCCTGCGATTAAATTCCTTGATTCGTGAGTCAGCCAG 498

Qy 167 GluHisGluAenIleGlyAspIleSerLysSerLeuValGlyLeuAenThrVal 186
 Db 499 GAACATGGGAATTCACCTGGATGCAATCTCCAAGTCTCTGCAGAGTCTGAATATGACACTG 558

Qy 187 LeuAspLeuGlnPheSerIleGluThrLeuAenGlyArgValGlnGlnAenAlaPheLys 206
 Db 559 CTTGATGTTCAACTCCATACAGAAACACTGAATGTGACAGTCCGTGTAATCTACAGCAAG 618

Qy 207 GlnGlnGluGluMetArgLysLeuGluGluArgLyleTyrAenAlaSerAlaGluLileLys 226
 Db 619 CAACAGGAGGACATCAGTAAATTTGGAGAACCTGTGTACAAAGTATACGAGAGAGTCAG 678

Qy 227 SerLeuAspGluLysGlnValTyrLeuGluGlnGluLileLysGlyGluMetLysLeuLeu 246
 Db 679 TCTGTGAAGAGAAACAAAGCCGACGTCAGCAAGGAGTAAACAGGAGAGTGAAGTATTG 738

Qy 247 AenAenIleThrAenAspLeuArgLysAspTrpGluHisSerGlnThrLeuLysAen 266
 Db 739 AACACATCACCACAGCTCAGACTGAAGAGTGGGAACACTCAGACACTGAAAAAC 798

Qy 267 IleThrLeuLeuGlnGly 272
 Db 799 ATCACCCTTCATTCAGGG 816

RESULT 10
 AK089178 1509 bp mRNA linear HTC 03-APR-2004
 LOCUS Mus musculus NOD-derived CD11c +ve dendritic cells cDNA, RIKEN
 DEFINITION full-length enriched library, clone:IF630005M03 product:macrophage
 scavenger receptor 1, full insert sequence.
 ACCESSION AK089178
 VERSION AK089178.1 GI:26354301
 KEYWORDS HTC; CAP trapper.

| | | | |
|----------------------|---|--------------------|--|
| SOURCE ORGANISM | Mus musculus (house mouse) | FEATURES source | Location/Qualifiers |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus. | | 1. .1509 /organism="Mus musculus" /mol_type="mRNA" /strain="NOD" /db_xref="PANTOM.DB.F630005M03" /db_xref="taxon:10090" /clone="F630005M03" /cell_type="NOD-derived CD11c +ve dendritic cells" /clone_lib="RIKEN full-length enriched mouse cDNA library" 53. .1117 /note="unnamed protein product; macrophage scavenger receptor 1 (MCD) MGI:98257, GB NM_031195, evidence: BLASTN, 99%, match=1461) putative" /codon_start=1 /protein_id="BAC40779.1" /db_xref="GI:26354302" /translation="MTKENTENQRLCPHEREDADCSSESVKFDPARSMTASLPHSTQNG PSVOKUKSPAAALIALYLLVFAVLIPVGVITVQALLNWMKGLVSRNTSDTSQGP MEKENTSNVEMRFTIMAHMKDMEERIQSISNKADLIDTGRFQNFNSATQORLNDIL LOKNSLILSVQEHGNSLDAISKLSQSLNMTLLDVQLHTETLHVRVRESATQQEDISK LEBRVYKVSARVQSVKEQAEVQEVKRVLNINITNDRKDKWHSOTUKNTFTFIO GPPGQCKGDRGLTGOTGPPGARGIRGIPGVKGRGQIGPPGGRNGPAGPKGPRSG SPGPKGQKSGSVGSGRSV" |
| REFERENCE AUTHORS | Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. | | polyA_signal 1489. .1494 /note="putative" 1509 /note="putative" |
| JOURNAL PUBMED | High-efficiency full-length cDNA cloning Meth. Enzymol. 303, 19-44 (1999) | | |
| REFERENCE AUTHORS | 10349636 | | |
| JOURNAL PUBMED | 2 | | |
| REFERENCE AUTHORS | Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuina, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. | | |
| JOURNAL PUBMED | RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer | | |
| REFERENCE AUTHORS | Genome Res. 10 (11), 1757-1771 (2000) | | |
| JOURNAL PUBMED | 4 | | |
| REFERENCE AUTHORS | The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium. | | |
| JOURNAL PUBMED | Functional annotation of a full-length mouse cDNA collection | | |
| REFERENCE AUTHORS | Nature 409, 685-690 (2001) | | |
| JOURNAL PUBMED | 5 | | |
| REFERENCE AUTHORS | The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team. | | |
| JOURNAL PUBMED | Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs | | |
| REFERENCE AUTHORS | Nature 420, 563-573 (2002) | | |
| JOURNAL PUBMED | 6 (bases 1 to 1509) | | |
| REFERENCE AUTHORS | Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Haneigaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirao, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ono, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y. | | |
| JOURNAL PUBMED | Direct Submission | | |
| REFERENCE AUTHORS | Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216) | | |
| JOURNAL PUBMED | cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. | | |
| REFERENCE AUTHORS | Tissues were provided by Dr. John Todd (Dept. of Medical Genetics Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome Trust/MRC building Addenbrookes Hospital Cambridge) whose assistance we gratefully acknowledge. Please visit our web site for further details. URL: http://genome.gsc.riken.jp/ URL: http://fantom.gsc.riken.jp/. | | |
| JOURNAL PUBMED | Alignment Scores: | | |
| REFERENCE AUTHORS | Pred. No.: 1.94e-82 Length: 1509 | | |
| JOURNAL PUBMED | Score: 801.00 Matches: 158 | | |
| REFERENCE AUTHORS | Percent Similarity: 76.7% Conservative: 46 | | |
| JOURNAL PUBMED | Best Local Similarity: 59.4% Mismatches: 60 | | |
| REFERENCE AUTHORS | Query Match: 38.9% Indels: 2 | | |
| JOURNAL PUBMED | DB: 4 Gaps: 2 | | |
| REFERENCE AUTHORS | US-10-618-570-2 (1-400) x AK089178 (1-1509) | | |
| JOURNAL PUBMED | Qy 8 ProbaspGlnGluAaspThrAspSerCysThrGluSerVallysPheAspAlaArgSer 27 | | |
| REFERENCE AUTHORS | Db 89 CCTCATGACGAGGAGGAGTGTGCTGCGAGTTTCAGAAATCCGGAATTTTGCACGCGTTC 148 | | |
| JOURNAL PUBMED | Qy 28 ValThrAlaLeuLeuProHisProIleAsnGlyProThrLeuGlnGluArgMetLys 47 | | |
| REFERENCE AUTHORS | Db 149 ATGACAGCATCCCTTCTCTCAGCAGCACTAAATGGCCCTCCGTTCCAGGAGATTGAAG 208 | | |
| JOURNAL PUBMED | Qy 48 SerTyrLysThrAlaLeuIleThrLeuTyrLeuIleValPheValValLeuValProIle 67 | | |
| REFERENCE AUTHORS | Db 209 TCCTTCAAGGTCGCCCTCATTCCTCTACCTCTGTGTGTTCAGTACTAATACCTGTT 268 | | |
| JOURNAL PUBMED | Qy 68 IleGlyIleValAlaAlaGlnLeuLeuLysTrpGluThrLysAsnCysThrValGlySer 87 | | |
| REFERENCE AUTHORS | Db 269 GTTGGAAATAGTGCAGCTGAGCTTTCGAAATGGGAAATAGAAAGCTGCTTAGTTTGTTC 328 | | |
| JOURNAL PUBMED | Qy 88 ValAsn--AlaAspIleSerProSerProGluGlyLysGlyAsnGlySerGluAspGlu 106 | | |
| REFERENCE AUTHORS | Db 329 CGTAAACAAGTATACATCTCAAGGTCCTATGGAAAGAAATACCAAGTAACGTGGAA 388 | | |
| JOURNAL PUBMED | Qy 107 MetArgPheArgGluAlaValMetGluArgMetSerAsnMetGluSerArgIleGlnTyr 126 | | |
| REFERENCE AUTHORS | Db 389 ATGAGATTT--ACAATTATCATGGCACATGAAGGACATGGAGGAGAGAGATCCAAAGC 445 | | |
| JOURNAL PUBMED | Qy 127 LeuSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsnPheGlnAsnPheSerIleThr 146 | | |
| REFERENCE AUTHORS | Db 446 ATTTCAAACTCAAAGCGGACCTTATAGACACGCGGCGCTTCCAGAAATTTCCAGCATGCA 505 | | |
| JOURNAL PUBMED | Qy 147 ThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsnSerLeuSerSerIleGln 166 | | |
| REFERENCE AUTHORS | Db 506 ACTGACCAAGACCTTAATGATATTTCTTCTGCGATTAATCTTCITGATTTTGTGAGTCG 565 | | |

Qy 167 GluHisGluAenIleGlyAspIleSerLysSerLeuValGlyLeuAsnThrThrVal 186
 Db 566 GAACATGGGAATTCTACTGGATGCAATCTCCAAAGTCCTTGCAGAGTCTGAATATGACACTG 625
 Qy 187 LeuAspLeuGlnPheSerIleGluThrLeuAenGlyArgValGlnGlnGluAenAlaPheLys 206
 Db 626 CTTGATGTTCAACTCCATACAGAACTGCGATGTCAGAGTCCGTGAATCTACAGCAAG 685
 Qy 207 GlnGlnGluGluMetArgLysLeuGluArgIleTyrAsnAlaSerAlaGluIleLys 226
 Db 686 CAACAGGAGGACATCAGTAAATTCGAGGAACGTGTGTACAAAGTATCATCAGCAGAGTCCAG 745
 Qy 227 SerLeuAspGluLysGlnValTyrLeuGlnGlnGluIleLysGlyGluMetLysLeuLeu 246
 Db 746 TCTGTGAAGAAGAAACAAGCCGACCTGGAAACAGGAAGTAAACAGGAAGTGGAGTATTG 805
 Qy 247 AsnAenIleThrAsnAspLeuArgLysLeuAspTyrGluHisSerGlnThrLeuLysAsn 266
 Db 806 AACACATCATCCACACGACTCAGACTGAAGACTGGGAACACTCACAGACACTGAAAAAC 865
 Qy 267 IleThrLeuLeuGlnGly 272
 Db 866 ATCACCTTCATTCAAGGG 883

RESULT 11
 BW956649
 LOCUS
 DEFINITION BW956649 full-length enriched swine cDNA library, adult alveolar macrophage Sus scrofa cDNA clone AMP010022D07 5', mRNA sequence.
 ACCESSION BW956649
 VERSION BW956649.1 GI:71974090
 KEYWORDS EST.
 SOURCE Sus scrofa (pig)
 ORGANISM Sus scrofa
 Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae; Sub.
 1 (bases 1 to 677)
 Uenishi,H., Eguchi,T., Suzuki,K., Sawazaki,T., Toki,D., Shinkai,H., Okumura,N., Hamasima,N. and Awata,T.
 TITLE PEDE (Pig EST Data Explorer): construction of a database for ESTs derived from porcine full-length cDNA libraries
 JOURNAL Nucleic Acids Res. 32 (1), D484-D488 (2004)
 PUBMED 14681463
 COMMENT Contact: Hirohide Uenishi
 Animal Genome Laboratory, Genome Research Department
 National Institute of Agrobiological Sciences
 2 Ikenodai, Tsukuba, Ibaraki 305-8602, Japan
 Tel: +81-29-838-8627
 Fax: +81-29-838-8627
 Email: huenishi@affrc.go.jp
 EST project with full-length enriched cDNA libraries carried out in Animal Genome Research Program (Japan) by National Institute of Agrobiological Sciences and STAFF-Institute
 Single pass sequencing of clones derived from oligo-capped cDNA library
 Vector sequences were eliminated by RepeatMasker version 2002/07/13 and crossmatch version 0.990319
 Low quality bases were trimmed based on the quality values.

FEATURES
 source
 1..677
 /organism="Sus scrofa"
 /mol_type="mRNA"
 /db_xref="taxon:9823"
 /clone="AMP010022D07"
 /tissue_type="alveolar macrophage"
 /dev_stage="adult"
 /clone_lib="full-length enriched swine cDNA library, adult alveolar macrophage"

ORIGIN

Alignment Scores:
 Pred. No.:

7.46e-82 Length: 677

Score: 791.50 Matches: 152
 Percent Similarity: 84.4% Conservative: 26
 Best Local Similarity: 72.0% Mismatches: 26
 Query Match: 38.4% Indels: 7
 DB: 5 Gaps: 1
 US-10-618-570-2 (1-400) x BW956649 (1-677)
 Qy 1 MetAlaGlnTrpAspPheProAspGlnGlnGlnAspThrAspSerCysThrGluSer 20
 Db 59 ATGGAACAGTGGGCATAGCTTCTCGATGATCAACCGGAGGACACTGACAGTTGATGGAATCT 118
 Qy 21 VallysPheAspAlaArgSerValThrAlaLeuLeuProProHisProLysAsnGlyPro 40
 Db 119 GTCAAGTTCGATGCTCGCTCAATGACAGCTTGTCTCTCAATCTCTAAATATGGCCCA 178
 Qy 41 ThrLeuGlnGluArgMetLysSerTyrLysThrAlaLeuIleThrLeuTyrLeuIleVal 60
 Db 179 ACGCTTCAAGAGAAAATGAAGTCTTCAAAAGCTGCACTGATTGCCCTTTATCTCTCTGTG 238
 Qy 61 PheValValLeuValProIleIleGlyIleValAlaAlaGlnLeuLeuLysTyrGluThr 80
 Db 239 TTTGTCTGCTGGTGCCTATCATCGAGTACTGCGAGCTCATCTCTGAAATGGGAAATG 298
 Qy 81 LysAenCysThrValGlySerValAsnAlaAspIleSerProSerProGluGlyLysGly 100
 Db 299 AAGAAATTGTCAGTTGGCTCAATTAGTGCA-----GGAAAGGA 337
 Qy 101 AsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsnMet 120
 Db 338 AATGACAGTGAATGACATGAGATTTTCGTGACGCTGTACAGAACATATGAGCAACATG 397
 Qy 121 GluSerArgIleGlnTyrLeuSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsnPhe 140
 Db 398 GAGAAGAGAAATTTAGTATCTTTTCAGATAAATGAAGCCCAATCTCTGAGAATCCGAGCATTT 457
 Qy 141 GlnAenPheSerIleThrThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsnSer 160
 Db 458 CAAATTTTCACTGCAATGACTGATCAAGATTTTAATGATGCTTTTCCAGCTAAATAAC 517
 Qy 161 LeuLeuSerSerIleGlnGluHisGluAenIleLeuGlyAspIleSerLysSerLeuVal 180
 Db 518 TTGCTTTCTTCATTACGGAACATGGAAATGTAATAATGAATCTCTAAGTCATTATA 577
 Qy 181 GlyLeuAenThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArgVal 200
 Db 578 AGTCTGAACACCACTTGTGATTTGAGTTTGCAGTCAATGTCGAACACTGAAAGGCAAGATC 637
 Qy 201 GlnGluAenAlaPheLysGlnGlnGluMet 211
 Db 638 CAAGAGAATACATTTAAACAACAAGAGAGATG 670

RESULT 12

DR004903

LOCUS

DEFINITION

DR004903

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

DR004903 768 bp mRNA linear EST 17-MAY-2005
 TCl09459 Human placenta, large insert, pCMV expression library Homo sapiens cDNA clone TCl09459 5' similar to Homo sapiens macrophage scavenger receptor 1 (MSR1), transcript variant SK-A11, mRNA sequence.

DR004903

DR004903.1

GI:66264776

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

1 (bases 1 to 768)

Birkett,C., Cho,J., Gau,Y., Hamer,R., Kelly,S., Kovacs,K., Liu,L.,

Liu,X., Porter,J., Sachs,A., Shu,Y., Sun,Z., Wong,J., Wu,M.,

Zhang,X., Jay,G. and He,W.

High-throughput cloning of full-length human cDNAs directly from

cDNA libraries optimized for large and rare transcripts

JOURNAL
COMMENT

Unpublished (2005)

Contact: Kovacs, KF
High throughput cDNA Cloning
Origene Technologies, Inc. (www.origene.com)
6 Taft Court, Suite 100, Rockville, MD 20850, USA
Tel: 301 340 3188
Fax: 301 340 8606
Email: cDNA@origene.com

This EST submission is part of an on-going human full-length
cloning project at Origene Technologies, Inc.

Please contact Origene for access.

Origene Technologies, Inc.

6 Taft Ct. Suite 100

Rockville, MD 20850

Tel: (301) 340-3188

http://www.origene.com

Seq primer: pCMV6 5prime forward vector primer, OriGene

Technologies Inc.

FEATURES
source

Location/Qualifiers
1..768
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="TC109459"
/tissue_type="Placenta"
/clone_lib="Human placenta, large insert, pCMV expression
library"

/note="Vector: pCMV6-XL4; Site 1: EcoRI; Site 2: XhoI/Sall
compatible end ligation; Oligo-dT primed reverse
transcription optimized for large and GC rich mRNA
transcripts, cDNA size selection, optimized ligation for
large inserts into mammalian expression vector, random
clones selected for end sequence verification of
full-length genes"

ORIGIN

Alignment Scores:

Pred. No.: 4.54e-79 Length: 768
Score: 768.50 Matches: 146
Percent Similarity: 84.4% Conservative: 32
Best Local Similarity: 69.2% Mismatches: 32
Query Match: 37.3% Indels: 1
DB: 8 Gaps: 1

US-10-618-570-2 (1-400) x DR004903 (1-768)

Qy 1 MetAlaGlnTTPAspAspPheProaspGlnGlnuaspThrAspSerCysThrGluSer 20
Db 123 ATGGAGCAGTGGGATCACTTTTCACATCAACAGGAGGACCTGATAGCTGCTCGAATCT 182
Qy 21 ValLysPheAspAlaArgSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
Db 193 GTGAATTTGATGCTCGCTCAATGACAGCTTGCTTCTCCGAATCTTAAACACAGCCT 242
Qy 41 ThrLeuGlnGluArgMetLysSerTyrLysThrAlaLeulleThrLeuTyrLeulleVal 60
Db 243 TCCCTTCAAGAGAACTCAAGTCCCTTCAAGTCTCAAGTCTCAAGTCTTACCTCTCTCGT 302
Qy 61 PheValValLeuValProIleGlyLeuValAlaGlnLeuLeuLysTyrGluThr 80
Db 303 TTTGAGTTCTCATCCCTCTCATTTGGAATAGTGGCAGCTCAACTCTCAAGTGGAAACG 362
Qy 81 LysAsnCysThrValGlySerValAsnAla---AspIleSerProSerProGluGlyLys 99
Db 363 AAGATTCTCTAGTTAGTTCACTAATGCAATGATATTAATCAAGTCTCACGGGAAA 422
Qy 100 GlyAsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsn 119
Db 423 GGAATGACAGCGAAGAGAAATGAGATTTCAAGAAGCTTTTATGGAACACATGAGCAAC 482
Qy 120 MetGluSerArgIleGlnTyrLeuSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsn 139
Db 483 ATGGAGAGAGAAATCCAGCATATTTTAGACATGGAAGCCAACTCTCATGGACACAGCAT 542

Qy 140 PheGlnAsnPheSerIleThrThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsn 159
Db 543 TTCCAAAATTTTCAGCATGACAACTGATCAAAAGATTTAATGACATTTCTCTCAGCTAAGT 602
Qy 160 SerLeuLeuSerSerIleGlnGluHisGluAsnIleIleGlyAspIleSerLysSerLeu 179
Db 603 ACCTTGTTTCTCAGTCCAGGACATGGGAATGCAATAGATGAATCTCAAGTCCTTA 662
Qy 180 ValGlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArg 199
Db 663 ATAAGTTTGAATACCACTTCTGATTTGAGCTCAACATAGAAAATCTCAATGGCAA 722
Qy 200 ValGlnGluAsnAlaPheLysGlnGlnGlu 210
Db 723 ATCAAGAGAGATACCTTCAAAACACAGAGAAA 755

RESULT 13

CB469554 626 bp mRNA linear EST 26-MAR-2003
LOCUS CB469554
DEFINITION sn09.A05.f sn Sus scrofa cDNA 5', mRNA sequence.
ACCESSION CB469554
VERSION CB469554.1 GI:29275939
KEYWORDS EST.
SOURCE Sus scrofa (pig)
ORGANISM Sus scrofa
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
Sus.

REFERENCE

1 (bases 1 to 626)

Neilan, J.G., Kutish, G.F., Lu, Z., Zsak, A. and Rock, D.L.
Sequence analysis of African swine fever virus infected and
non-infected porcine macrophage cDNA libraries

JOURNAL

COMMENT

Contact: Neilan JG

Plum Island Animal Disease Center

US Department of Agriculture, Agricultural Research Service

PO Box 848, Greenport, NY 11944-848, USA

Tel: 631 323 3133

Fax: 631 323 3044

Email: jneilan@piadc.ars.usda.gov

Single pass sequencing. Bases called with phred v0.020425.c and
trimmed with the aid of the trim.alt option. Vector identified by
cross match v0.990329 and lucy vl.17p.

Seq primer: M13 Forward.

FEATURES

source

Location/Qualifiers
1..626
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
/tissue_type="lymphoid"
/cell_type="macrophage"
/lab_hosts="DH10B"
/clone_lib="sn"

/note="Vector: pSPORT1; Site 1: NotI; Site 2: Sall;

Library made from pools of polyA selected RNA, isolated at
different times post-infection (0 to 16 hrs) from African
swine fever virus (isolate Pretoriuskop/96/4) infected
swine macrophages. Macrophages were derived from
peripheral blood mononuclear cells cultured for 48 hrs on
plastic in the presence of 30% L929 supernatant."

ORIGIN

Alignment Scores:

Pred. No.: 1.9e-75 Length: 626
Score: 736.50 Matches: 141
Percent Similarity: 84.7% Conservative: 25
Best Local Similarity: 71.9% Mismatches: 23
Query Match: 35.8% Indels: 7
DB: 6 Gaps: 1

US-10-618-570-2 (1-400) x CB469554 (1-626)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 27, 2006, 08:46:22 ; Search time 646 Seconds
(without alignments)
4126.745 Million cell updates/sec

Title: US-10-618-570-2

Perfect score: 2059

Sequence: 1 MAQWDFPDQEDTDSCTES.....DWKATRVGINIFTLRTQKE 400

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4996997 seqs, 332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/abs/ABSSWEB pool/US10618570/runat 27022006 063853 13038/app query fasta_1
-DB=N Geneseq QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOCPCL=0 -LOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=BLOSUM62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORES=ptc -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs805p
-USER=US10618570 @CGN 1 1 1096 @runat 27022006 063853 13038 -NCPU=6 -ICPU=3
-NO.MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N Geneseq 21:.*
1: geneseqn1980s.*
2: geneseqn1990s.*
3: geneseqn2000s.*
4: geneseqn2001as.*
5: geneseqn2001bs.*
6: geneseqn2002as.*
7: geneseqn2002bs.*
8: geneseqn2003as.*
9: geneseqn2003bs.*
10: geneseqn2003cs.*
11: geneseqn2003ds.*
12: geneseqn2004as.*
13: geneseqn2004bs.*
14: geneseqn2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|-------------|---------------------|
| 1 | 2059 | 100.0 | 5177 | 2 AAZ09997 | Aaz09997 Bovine sc |
| 2 | 1381.5 | 67.1 | 1588 | 2 AAZ04926 | Aaq04926 Sequence |
| 3 | 1381.5 | 67.1 | 1588 | 2 AAZ028540 | Aaq028540 DNA for b |
| 4 | 1381.5 | 67.1 | 1807 | 2 AAZ04925 | Aaq04925 Sequence |

| | | | | | | |
|----|--------|------|------|----|----------|------------|
| 5 | 1042.5 | 50.6 | 1077 | 10 | ADB89012 | Human DNA |
| 6 | 1042.5 | 50.6 | 1347 | 2 | AAQ21550 | Gene enco |
| 7 | 1042.5 | 50.6 | 1347 | 2 | AAQ47731 | Human sca |
| 8 | 1042.5 | 50.6 | 1347 | 4 | AAF29165 | Human DNA |
| 9 | 1042.5 | 50.6 | 1367 | 2 | AAQ43260 | Type II m |
| 10 | 1042.5 | 50.6 | 1367 | 10 | ACA56662 | Human sig |
| 11 | 1042.5 | 50.6 | 1367 | 12 | ADI56458 | Human pol |
| 12 | 1042.5 | 50.6 | 2028 | 2 | AAQ21549 | Gene enco |
| 13 | 1042.5 | 50.6 | 2028 | 2 | AAQ47730 | Human sca |
| 14 | 1042.5 | 50.6 | 2028 | 2 | AAQ43259 | Type I ma |
| 15 | 1042.5 | 50.6 | 2028 | 4 | AAF29164 | Human DNA |
| 16 | 1042.5 | 50.6 | 2028 | 12 | ADI62122 | Human inac |
| 17 | 1042.5 | 50.6 | 2877 | 13 | ADQ38907 | Human SNP |
| 18 | 1042.5 | 50.6 | 3336 | 13 | ADQ38905 | Human SNP |
| 19 | 1042.5 | 50.6 | 3525 | 13 | ADQ38904 | Human SNP |
| 20 | 1042.5 | 50.6 | 3704 | 6 | AAQ94868 | Human DNA |
| 21 | 1042.5 | 50.6 | 3719 | 13 | ADQ38906 | Human SNP |
| 22 | 1042.5 | 50.6 | 3791 | 13 | ADQ38903 | Human SNP |
| 23 | 1042.5 | 50.6 | 3898 | 13 | ADP24728 | PRO polyp |
| 24 | 1039.5 | 50.5 | 2037 | 2 | AAQ68796 | Macrophag |
| 25 | 1032.5 | 50.1 | 2037 | 2 | AAQ28539 | DNA for h |
| 26 | 994.5 | 48.3 | 1330 | 12 | ADG39681 | Human cDN |
| 27 | 994.5 | 48.3 | 1330 | 12 | ADG39730 | Human cDN |
| 28 | 688 | 33.4 | 459 | 12 | ADP86098 | Barley al |
| 29 | 688 | 33.4 | 484 | 2 | AAQ68502 | Barley al |
| 30 | 680.5 | 33.1 | 1522 | 12 | ADM97492 | CD1d-IgG- |
| 31 | 678.5 | 33.0 | 1746 | 14 | AEA28083 | Circularl |
| 32 | 678 | 32.9 | 387 | 14 | ADV78150 | Chicken a |
| 33 | 678 | 32.9 | 459 | 2 | AAQ80197 | Chicken a |
| 34 | 678 | 32.9 | 480 | 14 | ADV78151 | Barley al |
| 35 | 678 | 32.9 | 486 | 3 | AAZ49865 | Potato pr |
| 36 | 678 | 32.9 | 576 | 3 | AAZ49855 | pGEMav av |
| 37 | 678 | 32.9 | 604 | 2 | AAQ73194 | Wild-type |
| 38 | 678 | 32.9 | 604 | 12 | ADP26979 | Chicken a |
| 39 | 678 | 32.9 | 823 | 12 | ADM97488 | CD1d-IgG- |
| 40 | 678 | 32.9 | 871 | 12 | ADM97490 | CD1d-IgG- |
| 41 | 677.5 | 32.9 | 873 | 14 | AEA28087 | Circularl |
| 42 | 670 | 32.5 | 540 | 14 | ADX98285 | Chicken a |
| 43 | 670 | 32.5 | 540 | 14 | ADX98286 | Chicken a |
| 44 | 647 | 31.4 | 401 | 12 | ADQ85120 | DNA seque |
| 45 | 612 | 29.7 | 433 | 12 | ADQ85126 | DNA seque |

ALIGNMENTS

| | |
|----------|--|
| RESULT 1 | |
| AAZ09997 | |
| ID | AAZ09997 standard; cDNA; 5177 BP. |
| XX | |
| AC | AAZ09997; |
| XX | |
| DT | 03-DEC-1999 (first entry) |
| XX | |
| DE | Bovine scavenger receptor class A (ScR)/avidin fusion protein cDNA. |
| XX | |
| KW | Scavenger receptor class A; ScR; avidin; fusion protein; bovine; ECD; |
| KW | membrane-spanning domain; extracellular domain; biotin-binding activity; |
| KW | endocytosis; ss. |
| XX | |
| OS | Synthetic. |
| OS | Bos taurus. |
| XX | |
| FT | Key Location/Qualifiers |
| CDS | 1071..2269 |
| FT | /*tag= a |
| FT | /product= "ScR/avidin fusion protein" |
| XX | |
| PN | W09942577-A2. |
| XX | |
| PD | 26-AUG-1999. |
| XX | |
| PF | 23-FEB-1999; 99WO-GB0000546. |
| XX | |

PR 23-FEB-1998; 98GB-00003757.
 PR 24-JUN-1998; 98GB-00013653.
 XX (EURO-) EUROGENE LTD.
 XX Yla-Herttua S, Kulomaa M, Lehtolainen P, Marjomaki V, Airenne K;
 FI MPI; 1999-561345/47.
 DR P-PSDB; AAY03098.
 XX New fusion proteins having an extracellular domain with biotin-binding
 PT activity, used to target biotinylated molecules to specific sites in
 PT tissues.
 XX Disclosure; Page 15-21; 23pp; English.
 XX This invention describes a novel protein (A) which comprises a membrane-
 CC spanning domain and an extracellular domain (ECD), where the ECD
 CC comprises biotin-binding activity. Using the proteins or encoding nucleic
 CC acid molecules it is possible to target biotinylated molecules to
 CC specific sites in tissues. Molecules targeted in this way may be taken up
 CC by the tissues or cells by endocytosis, allowing the molecules to exert
 CC their effects within or on the cell. This sequence encodes a bovine
 CC scavenger receptor class A/avidin fusion protein which is used in the
 CC description of the invention
 XX
 SQ Sequence 5177 BP; 1356 A; 1222 C; 1390 G; 1209 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 3,86e-177 Length: 5177
 Score: 2059.00 Matches: 400
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 2 Gaps: 0
 US-10-618-570-2 (1-400) x AAZ09997 (1-5177)
 QY 1 MetAlaGlnTrpAspAspPheProAspGlnGlnGluAspThrAspSerCysThrGluSer 20
 DB 1071 ATGGCACAGTGGGATGACTTTCTCTGATCAGCAAGAGGACACTGCAGCTGTACAGAGTCT 1130
 QY 21 ValLysPheAspAlaArgSerValThrAlaLeuLeuProProHisProLysAsnGlyPro 40
 DB 1131 GTGAAGTTCGATGCTCGCTCAGTGACAGCTTTGCTCTCCCTCCCATCTCTAAANAATGGGCCA 1190
 QY 41 ThrLeuGlnGluArgMetLysSerTyrLysThrAlaLeuLeuThrLeuLysLeuVal 60
 DB 1191 ACTCTTCAAGAGAGGATGAAGTCTTATAAACTGCACCTGATCACCTTTATCTCATTTGTG 1250
 QY 61 PheValValLeuValProIleIleGlyIleValAlaGlnLeuLeuLysTrpGluThr 80
 DB 1251 TTTGTAGTTCCTGTCGCCATCATTTGGCATATGTCAGCTCAGCTCCTGAAATGGGAAACG 1310
 QY 81 LysAsnCysThrValGlySerValAsnAlaAspIleSerProSerProGluGlyLysGly 100
 DB 1311 AAGAATTCACGGTTGGCTCAGTATATGCAATATATCTCCAGTCCGGAAGGCAAGGA 1370
 QY 101 AsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsnMet 120
 DB 1371 AATGCACTGAAGATGAATGAGATTTTCGAGAGCTGTGTGGAACGCATGAGCAACATG 1430
 QY 121 GluSerArgIleGlnTyrLeuSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsnPhe 140
 DB 1431 GAAAGCAGAAATCCAGTATCTTTTCAGATAATGAAGCCAATCTCCTAGATGCTTAAGAATTC 1490
 QY 141 GlnAsnPheSerIleThrThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsnSer 160
 DB 1491 CAANAATTCAGATAAACAACATGATCAAGATTTTATGATGTTCTTTTCCAGCTAAATTC 1550
 QY 161 LeuLeuSerSerIleGlnGluHisGlnAsnIleIleGlyAspIleSerLysSerLeuVal 180
 DB 1551 TTACTTTCTCCATCCAGGAACATGAGAAATATCATAGGGGATATCTCCCAAGTCTTAGTA 1610

QY 181 GlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArgVal 200
 DB 1611 GGTCTGAACACCACAGTACTTTGATTTTCAGTTTCAGTATTGAAACACTGAATGGCAGAGTC 1670
 QY 201 GlnGluAsnAlaPheLysGlnGlnGluMetArgLysLeuGluGluArgIleTyrAsn 220
 DB 1671 CAAGAGAAATGCATTTTAAACAACAAGAGAGATGCGTAAATTTAGAGGAGGTATATACAAT 1730
 QY 221 AlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyrLeuGluGlnGluIleLys 240
 DB 1731 GCATCAGCAGAAATTAAGTCTCTAGATGAAACCAAGTATATTTTGGAAACAGGAATAAAA 1790
 QY 241 GlyGluMetLysLeuLeuAsnAsnIleThrAsnAspLeuArgLeuLysAspTrpGluHis 260
 DB 1791 GGGGAAATGAACACTGTTGAATAATATCATTAATGCTGAGGCTGAAGAGTTGGGAACAT 1850
 QY 261 SerGlnThrLeuLysAsnIleThrLeuLeuGlnGlyAlaArgLysCysSerLeuThrGly 280
 DB 1851 TCTCAGACATTTGAAATAATATCATCTTACTCAAGTGCAGAGAAAGTGTCTGCTCAGCTGGG 1910
 QY 281 LysTrpThrAsnAspLeuGlySerAsnMetThrIleGlyAlaValAsnSerArgGlyGlu 300
 DB 1911 AAATGGACCAACGATCTGGGCTCCAAACATGACCATCGGGGCTGTGAAACAGCAGAGGTGAA 1970
 QY 301 PheThrGlyThrTyrIleThrAlaValThrAlaThrSerAsnGluIleLysGluSerPro 320
 DB 1971 TTCACAGGCACCTACATCAGAGCGGTAAACGCCACATCAATAGATCAAAAGATCACCA 2030
 QY 321 LeuHisGlyThrGlnAsnThrIleAsnLysArgThrGlnProThrPheGlyPheThrVal 340
 DB 2031 CTGCATGGGACACAAACACCATCAACAAGAGGACCCAGCCACCTTTGGCTTCCACCGTC 2090
 QY 341 AsnTrpLysPheSerGluSerThrThrValPheThrGlyGlnCysPheIleAspArgAsn 360
 DB 2091 AATTGGAAATTTTCAGAGTCCACCACCTGCTTTCACGGGCCAGTGTCTTCATAGACAGGAAT 2150
 QY 361 GlyLysGluValLeuLysThrMetTrpLeuLeuArgSerValAsnAspIleGlyAsp 380
 DB 2151 GGGAAAGAGGTGCTTGAAAGACCATGTGGCTGCTGGGTCAAAGTGTATATGACATTTGGTAT 2210
 QY 381 AspTrpLysAlaThrArgValGlyIleAsnIlePheThrArgLeuArgThrGlnLysGlu 400
 DB 2211 GACTGGAAAGCTACCAGGTTCGGCATCAACATCTTCACTCGCTGCGCACACACAGAGGAG 2270
 RESULT 2
 AAQ04926
 ID AAQ04926 standard; DNA; 1588 BP.
 XX
 AC AAQ04926;
 XX
 DT 25-MAR-2003 (revised)
 DT 31-OCT-2002 (revised)
 DT 23-OCT-1990 (first entry)
 XX
 DE Sequence encoding portion of scavenger receptor protein with affinity for
 DE acylated low density lipoprotein (aLDL).
 XX
 XX Acylated low density lipoprotein; aLDL; atherosclerotic plaque; ds.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT CDS 4..1362
 FT /*tag= a
 XX
 XX W09005748-A.
 XX
 PD 31-MAY-1990.
 XX
 PF 15-NOV-1988; 88US-00272002.
 XX
 XX 15-NOV-1988; 88US-00272002..

PR 09-AUG-1989; 89US-00391486.
 XX (MASI) MASSACHUSETTS INST TECHNOLOGY.
 PA
 PI Kodama T, Krieger M;
 XX
 DR WPI; 1990-193408/25.
 DR P-PSDB; AAR05509.
 XX
 PT New receptor protein - has affinity for acetylated low density
 PT lipoprotein and corresponding antibodies and DNA sequences.
 XX
 PS Disclosure; Page ?; 79pp; English.
 XX
 CC Receptor protein, and fragments and analogues thereof may be immobilised
 CC on a support and used in assay and purification of the aLDL target.
 CC Labelled Abs, raised to the protein may be injected into the vascular
 CC system to detect the presence of atherosclerotic plaques. (Updated on 31-
 CC OCT-2002 to add missing OS field.) (Updated on 25-MAR-2003 to correct PR
 CC field.) (Updated on 25-MAR-2003 to correct PA field.) (Updated on 25-MAR-
 CC 2003 to correct PI field.)
 XX
 SQ Sequence 1588 BP; 506 A; 304 C; 367 G; 411 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 4.78e-116 Length: 1588
 Score: 1381.50 Matches: 307
 Percent Similarity: 79.4% Conservative: 24
 Best Local Similarity: 73.6% Mismatches: 66
 Query Match: 67.1% Indels: 22
 DB: 2 Gaps: 6

US-10-618-570-2 (1-400) x AAQ04926 (1-1588)

Qy 1 MetAlaGlnTTPAspAspPheProAspGlnGlnuAspThrAspSerCysThrGluSer 20
 Db 4 ATGGCACAGTGGGATGACTTTCCCTGATCAGCAAGAGGACTGACAGCTGTACAGATCT 63
 Qy 21 ValLysPheAspAlaArgSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
 Db 64 GTGAAGTTGATGCTGCTCAGTCAGCAGCTTTGCTTCTCCCTCCATCTTAAATGGCCCA 123
 Qy 41 ThrLeuGlnGluArgMetLysSerTyrLysThrAlaLeuLeuThrLeuTyrLeuVal 60
 Db 124 ACTCTCAAGAGAGGATGAGTCTTATAAACTGCATCATCATCATCATCATCATCATCAT 183
 Qy 61 PheValValLeuValProIleGlyLeuValAlaGlnLeuLeuLysTyrGluThr 80
 Db 184 TTTGTAGTTCTGTCGCCCATCATTTGGCATATGTCAGCTCAGCTCCTGAAATGGGAACG 243
 Qy 81 LysAsnCysThrValGlySerValAsnAlaAspIleSerProSerProGluGlyLysGly 100
 Db 244 AAGAAATTCACGGTGGCTCAGTTAATGAGATATATCTCCAGTCCGGAAGGCAAGGA 303
 Qy 101 AsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsnMet 120
 Db 304 AATGCGAGTGAAGATGAATGAGATTTCCAGAAGCTGTGATGGAACGATGAGCAACATG 363
 Qy 121 GluSerArgIleGlnTyrLeuSerAspAsnGluAlaLeuLeuLeuAspAlaLysAsnPhe 140
 Db 364 GAAAGCAGAAATCCAGTATCTTTTCAGATATATGAGCCAAATCTCCTAGATGCTAAGATTC 423
 Qy 141 GlnAsnPheSerIleThrThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsnSer 160
 Db 424 CAAATTTACATACACTGATGATCAAGATTTATGATGTTCTTTCCAGCTAAATTC 483
 Qy 161 LeuLeuSerSerIleGlnGluHisGluAsnIleLeuGlyAspIleSerLysSerLeuVal 180
 Db 484 TTACTTTCTCCATCCAGGAACATGAGAAATATCATAGGGATATCTCCAAAGTCATTAGTA 543
 Qy 181 GlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArgVal 200
 Db 544 GGTCTGAACACACAGTACTTTGATTTGAGTTTCAGTTTCAGTATTTGAAACACTGAATGGCAGATC 603

Qy 201 GlnGluAsnAlaPheLysGlnGlnGluMetArgLysLeuGluGluArgIleTyrAsn 220
 Db 604 CAAGAGAAATGCAATTTAAACAACAAGAGAGATGCGTAAATTTAGAGGCGTATATACAAT 663
 Qy 221 AlaSerAlaGluIleLysSerLeuAspGlnLysGlnValTyrLeuGluGlnGluLys 240
 Db 664 GCATCAGCAGAAATTAAGTCTCTAGATGAAACAAAGTATATATTTGGAAACAGGAATAAAA 723
 Qy 241 GlyGluMetLysLeuLeuAsnAsnIleThrAsnAspLeuArgLeuLysAspTyrGluHis 260
 Db 724 GGGGAATGAACCTGTTGAATATATCACTAATGATCTGAGGCTGAAGGATGGGAACAT 783
 Qy 261 SerGlnThrLysAsnIleThrLeuLeuGlnGlyAla----- 273
 Db 784 TCTCAGACATTTGAAATAATATCACTTTACTCAAGGCTCTCTCGACCTCCAGGTTGAAAAA 843
 Qy 274 -----ArgLysCysSerLeuThrGlyLysThrAsnAspLeuGly---SerAsnMet 290
 Db 844 GGAGATAGAGGCC-TCCTGGACA---AAATGGTATACACAGGCTTCCAGGCTCTAATAGG 899
 Qy 291 ThrIleGlyAlaValAsnSerArgGlyGluPheThrGlyThrTyrIleThrAlaValThr 310
 Db 900 TACTCCAGGCTTTAAAGGTGATCGGGGATCTCTGTTTACTCGAGTTCGAGGATTCCTCC 959
 Qy 311 AlaThrSerAsnGluIleLysGluSer-ProLeuHisGlyThrGlnAsnThrIleAsnLys 330
 Db 960 AGGACCAATGGGAAGACCGGAGCAGGACTTAAATGGAACAAAAGGCCAGAGGGAGA 1019
 Qy 330 sArgThrGlnProThrPhe-----GlyPheThrVal-AsnTyrLysPheSerGlu 347
 Db 1020 AAAACCCACTCCACCATCCAAACACAATCTAATACAGTCCGACTCGGTGGTGGCAGCGG 1079
 Qy 347 erThrThrValPheThrGlyGlnCysPheIleAspArgAsnGlyLysGluValLeuLysT 367
 Db 1080 CCCTCAGAGGCGAGATGGAGATTTTTCAGGAAGCCAGTGGGTACCGTGTGTGACGA 1139
 Qy 367 hrMetTyrLeuLeuArgSerSerVal-AsnAspIleGlyAsp---AspTyrLysAlaThr 385
 Db 1140 CCGC-TGGGAACTCGTCGAGGAGTGTGTCGTCGACAGGAGCTTCGGATACAAAGTGTTC 1198
 Qy 386 ArgValGlyIleAsn-----IlePheThrArgLeuArgThrGln 398
 Db 1199 AAGTGTGCTAAGCAGGCTTATTTTGGAAAGAGTACGGGTCCAA 1243

RESULT 3
 AAQ28540
 ID AAQ28540 standard; DNA; 1588 BP.
 XX
 AC AAQ28540;
 XX
 DT 25-MAR-2003 (revised)
 DT 16-FEB-1993 (first entry)
 XX
 DE DNA for bovine sol. scavenger receptor.
 XX
 KW Macrophage; endotoxaemia; radiolabelled; toxic; degenerate; ss.
 XX
 OS Bos taurus.
 XX
 FH Key Location/Qualifiers
 CDS 4..1362.
 FT /*tag= a
 FT
 XX
 PN W09214482-A1.
 XX
 PD 03-SEP-1992.
 XX
 PF 21-FEB-1992; 92WO-US001370.
 XX
 PR 22-FEB-1991; 91US-00662227.
 XX
 PA (MASI) MASSACHUSETTS INST TECHNOLOGY.

PD 31-MAY-1990.
 XX 15-NOV-1988; 88US-00272002.
 XX 15-NOV-1988; 88US-00272002.
 PR 09-AUG-1989; 89US-00391486.
 XX (MASI) MASSACHUSETTS INST TECHNOLOGY.
 XX Kodama T, Krieger M;
 XX WPI; 1990-193408/25.
 DR P-PSDB; AAR05508.
 XX
 PT New receptor protein - has affinity for acetylated low density
 PT lipoprotein and corresponding antibodies and DNA sequences.
 XX
 XX Disclosure; Page ?; 79pp; English.
 PS
 CC Receptor protein, and fragments and analogues thereof may be immobilised
 CC on a support and used in assay and purification of the aLDL target.
 CC Labelled Abs, raised to the protein may be injected into the vascular
 CC system to detect the presence of atherosclerotic plaques. (Updated on 31-
 CC OCT-2002 to add missing OS field.) (Updated on 25-MAR-2003 to correct PR
 CC field.) (Updated on 25-MAR-2003 to correct PA field.) (Updated on 25-MAR-
 CC 2003 to correct PI field.)
 XX
 XX Sequence 1807 BP; 556 A; 384 C; 406 G; 461 T; 0 U; 0 Other;
 SQ
 Alignment Scores:
 Pred. No.: 5.63e-116 Length: 1807
 Score: 1381.50 Matches: 288
 Percent Similarity: 85.2% Conservative: 16
 Best Local Similarity: 80.7% Mismatches: 34
 Query Match: 67.1% Indels: 20
 Gaps: 4
 DB:
 US-10-618-570-2 (1-400) x AAQ04925 (1-1807)
 QY 1 MetAlaGlnTrpAspAspPheProAspGlnGluAspThrAspSerCysThrGluSer 20
 DB 36 ATGCACAGTGGGATGACTTTCCTGATCAGCAAGGACCTGACAGCTGTACAGATCT 95
 QY 21 ValLysPheAspAlaArgSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
 DB 96 GTGAAGTTCGATGCTCGCTCAGTCAGACGCTTGTCTCCCTCCCATCTTAAATGGCCCA 155
 QY 41 ThrLeuGlnGluArgMetLysSerTyrlsThrAlaLeuLeuLeuLeuLeuVal 60
 DB 156 ACTCTTCAAGAGAGGATGAAGTCTTATAAACTGCACCTGATCACCTTTATCTCATTTGTG 215
 QY 61 PheValValLeuValProIleIleGlyLeValAlaAlaGlnLeuLeuLysTrpGluThr 80
 DB 216 TTGTGTAGTTCGTGGCCCATCTTGGCATATGTCAGCTCAGCTCCCTGAAATGGGAACG 275
 QY 81 LysAsnCysThrValGlySerValAsnAlaAspIleSerProSerProGluGlyLysGly 100
 DB 276 AAGATTGCACGGTTGGCTCAGTTAATGCAGATATATCTCCAGTCCGGAAGGCAAGGA 335
 QY 101 AsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsnMet 120
 DB 336 AATGCGAGTGAAGATGAATGAGATTTTCGAGAAGCTGTGTGATGGAACGCGATGAGCAACATG 395
 QY 121 GluSerArgIleGlnTyrlsLeuSerAspAsnGluAlaLeuLeuLeuLeuLeuLeuPhe 140
 DB 396 GAAAGCAGATCCAGTATCTTTTCAGATAATGAAGCCAAATCTCTAGATGCTTAAGAAATTC 455
 QY 141 GlnAsnPheSerIleThrThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsnSer 160
 DB 456 CAAATTTTCAGCATPACAACTGATCAAGATTTAATGATGTCTTTTCCAGCTAAATTC 515
 QY 161 LeuLeuSerSerIleGlnGluHisGluAsnIleLeuGlyAspIleSerLysSerLeuVal 180

DB 516 TTACTTTCTCCATCCAGGAACATGAGAATATCATAGGGATATCTCCAAGTCATTAGTA 575
 QY 181 GlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArgVal 200
 DB 576 GGTCTGAACACCACTGACTTGTGAGTTTCAGTATTTGAAACACTGAATGGCAGAGTC 635
 QY 201 GlnGluAsnAlaPheLysGlnGlnGluMetArgLysLeuGluGluArgIleLysAsn 220
 DB 636 CAAGAGAAATGCATTTAAACCAACAGAGGAGATGCGTAAATTTAGAGGAGCGGTATATACAAT 695
 QY 221 AlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyrlsLeuGlnGlnGluLys 240
 DB 696 GCATCAGCAGAAATTAAGTCTCTAGATGAANAACNAGTATATTTGGACAGGAATATAAA 755
 QY 241 GlyGluMetLysLeuLeuAsnAsnIleThrAsnAspLeuArgLeuLysAspTrpGluHis 260
 DB 756 GGGGAATGAACACTGTTGAATAATATCACTAATGATCTGAGGCTGAAGGATTTGGGAACAT 815
 QY 261 SerGlnThrLysAsnIleThrLeuGlnGlyAla----- 273
 DB 816 TCTCAGACATTTGAAAAATATCACTTTACTCAAGGTCCTCTCGACCTCCAGGTGAAAAA 875
 QY 274 -----ArgLysCysSerLeuThrGlyLysTrpThrAsnAspLeuGly---SerAsnMet 290
 DB 876 GGAGATAGAGGCC-TCTCGACA---AATGTGTATACCAAGGCTTTCCAGGTCTAATAGG 931
 QY 291 ThrIleGlyAlaValAsnSerArgGlyGluPheThrGlyTrpIleThrAlaValThr 310
 DB 932 TACTCCAGGTCCTTAAAGGTGATCGGGGATCTCTGGTTTACCTGGAGTTCGAGGATTCCTCC 991
 QY 311 AlaThrSerAsnGluIleLysGluSerProLeuHisGlyThrGlnAsnThrIleAsnLys 330
 DB 992 AGGACCAATGGGAAGACCGGGAAGCCAGGACTTATTTGGCAAAAGCCAGAGAGGAGA 1051
 QY 331 ArgThrGlnProThrPheGlyPheThrValAsnTrpLysPheSerGluSer 347
 DB 1052 AAAAGGGAG-----TGAAGCATGCAAGACC 1078
 RESULT 5
 ADB89012
 ID ADB89012 standard; DNA; 1077 BP.
 XX
 AC ADB89012;
 XX
 DT 18-DEC-2003 (first entry)
 XX
 DE Human DNA encoding the macrophage scavenging receptor.
 XX
 KW Human; db; macrophage scavenger receptor; heptad repeat;
 KW reduced immunogenicity; trimerisation; CD40L; immune response;
 KW T-cell activation; isotype switching; gene.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..1077
 FT /*tag= a
 FT /product= "Macrophage scavenger receptor"
 XX
 FN US2003119149-A1.
 XX
 XX 26-JUN-2003.
 XX
 XX 20-DEC-2002; 2002US-00326186.
 XX
 XX 21-DEC-2001; 2001US-0343315P.
 XX
 XX (REDD/) REDDY P.
 XX
 XX Reddy P;
 XX
 XX WPI; 2003-863437/80.
 DR P-PSDB; ADB89003.

XX New fusion polypeptide that is capable of forming a trimer, useful in
PT preparing a composition for treating immune related or inflammatory
PT disorders.

XX Example; Fig 4; 24pp; English.

XX The present invention relates to a polypeptide comprising a fusion
CC protein capable of forming a trimer and comprising a heptad repeat
CC containing region of the human macrophage scavenging factor protein
CC appearing as ADB89007 fused to a second protein (especially a member of
CC the tumour necrosis superfamily, TNFSF e.g. CD40L). CD40L is involved in
CC isotype switching and is important in T-cell activation in the immune
CC response. The fusion protein has reduced immunogenicity due to its
CC ability to form trimers. Also included are the nucleic acids encoding the
CC fusion proteins, host cells comprising the nucleic acids (used to produce
CC the fusion proteins) and a recombinant vector comprising the nucleic acid.
CC The fusion proteins are used to deliver a therapeutic protein with
CC reduced immunogenicity to a patient in need of the protein. The present
CC sequence encodes the human macrophage scavenging factor protein. Note:
CC The authors have mixed up the descriptions of the sequences as they
CC appear in the figures, therefore there is a possibility that the
CC description assigned by the indexer may be wrong.

XX SQ Sequence 1077 BP; 364 A; 221 C; 242 G; 250 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2,29e-85 Length: 1077
Score: 1042.50 Matches: 197
Percent Similarity: 87.2% Conservative: 41
Best Local Similarity: 72.2% Mismatches: 34
Query Match: 50.6% Indels: 1
DB: 10 Gaps: 1

US-10-618-570-2 (1-400) x ADB89012 (1-1077)

QY 1 MetAlaGlnTrpAspAspPheProAspGlnGlnGluAspThrAspSerCysThrGluSer 20
DB 1 ATGGAGCAGTGGGATCACTTTTCACATCAACAGAGGACACTGATAGCTGCTCGAATCT 60
QY 21 ValLysPheAspAlaAargSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
DB 61 GTGAAATTTGATGCTCGCTCAATGACAGCTTTGCTTCCTCGAATCTTAAACACAGCCCT 120
QY 41 ThrLeuGlnGluAargMetLysSerTyrLysThrAlaLeuLeuThrLeuTyrLeuVal 60
DB 121 TCCCTTCAAGAGAAACTCAAGCTCTTCAAGCTGCACTGATGCGCTTTACCTCTCGTG 180
QY 61 PheValValLeuValProIleGlyIleValAlaAlaGlnLeuLeuTyrTrpGluThr 80
DB 181 TTTCAGTTCATCCCTCTCATTTGGAATAGTGGAGCTCAACTCCTGAAGTGGGAACG 240
QY 81 LysAsnCysThrValGlySerValAsnAla---AspIleSerProSerProGluGlyLys 99
DB 241 AAGAATTGCTCAGTTAGTTCAACTTAATGCAATATGATATACTCAAGTCTCACGGGAAA 300
QY 100 GlyAsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluAargMetSerAsn 119
DB 301 GGAATGACACGGAAGAGAAATGAGATTTTCAAGAAAGTCTTTATGGAACACATCAGCAAC 360
QY 120 MetGluSerArgIleGlnTyrLeuSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsn 139
DB 361 ATGGAGAAGAAATCCACATATTTAGCATGGAAGCCACCTCATGGACACAGAGCAT 420
QY 140 PheGlnAsnPheSerIleThrThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsn 159
DB 421 TTCCAAAATTTTCAGCATGACAACTGATCAAGATTTAATGACATTTCTTCGACCTAAGT 480
QY 160 SerLeuLeuSerSerIleGlnGluHisGluAsnIleGlyAspIleSerLysSerLeu 179
DB 481 ACCTTGTTTCTCAGTCCAGGACATGGGAATGCAATAGATGAATCTCCAACTCCTTA 540
QY 180 ValGlyLeuLeuAsnThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArg 199

DB 541 ATAAGTTTGAATACCACTTGCCTTGATTTCAGCTCAACATAGAAATCTGAATGGCAA 600
QY 200 ValGlnGluAsnAlaPheLysGlnGlnGlnMetArgLysLeuGluGluAargIleTyr 219
DB 601 ATCCAGAGAATACCTTCAAAACAACAAGAGAAATAGTAAATAGAGGACGCTGTTTAC 660
QY 220 AsnAlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyrLeuGluGluIle 239
DB 661 AATGTATTCACAGAAATATATGGCTATGAAAGAAACAAGTGCATTTGGAAACAGGAATA 720
QY 240 LysGlyGlnMetLysLeuLeuAsnAsnIleThrAsnAspLeuAargLeuLysAspTyrGlu 259
DB 721 AAGGAGAGAGTGAAGTACTGATAACATCACTAATGATCTCAGACTGAAAGATTGGAA 780
QY 260 HisSerGlnThrLeuLysAsnIleThrLeuLeuGlnGly 272
DB 781 CATTCTCAGACCTTGAGAAATATCACTTTAATTCAAGGT 819

RESULT 6

AAQ21550
ID AAQ21550 standard; DNA; 1347 BP.

XX AAQ21550;

XX 25-MAR-2003 (revised)
DT 19-MAY-1992 (first entry)

XX Gene encoding human scavenger receptor type II.

XX Lipoproteins; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS 47..1124

XX FT /*tag= a

XX FT /product= "scavenger_receptor_type_II"

XX JF03290184-A.

XX 19-DEC-1991.

XX 06-APR-1990; 90JP-00090274.

XX 06-APR-1990; 90JP-00090274.

XX (CHUS) CHUGAI PHARM CO LTD.

XX WPI; 1992-051436/07.

XX P-PSDB; AAR21513.

XX New scavenger receptor-producing animal cells - which have been
PT transformed with vector contg. gene for coding human scavenger receptor I
PT or II type under control of promoter, etc.

XX Disclosure; Fig 2; 9pp; Japanese.

XX The gene can be expressed from a cytomegalovirus promoter in a host cell
CC e.g. CHO cell. The resulting recombinant scavenger receptor can be used
CC to detect modified lipoproteins or modified substances in the blood. See
CC also AAQ21549. (Updated on 25-MAR-2003 to correct PA field.)

XX SQ Sequence 1347 BP; 420 A; 297 C; 298 G; 332 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 3,05e-85 Length: 1347
Score: 1042.50 Matches: 197
Percent Similarity: 87.2% Conservative: 41
Best Local Similarity: 72.2% Mismatches: 34
Query Match: 50.6% Indels: 1
DB: 2 Gaps: 1


```

Db 467 TTTCAAAATTTTCAGCATGACAACTGATCAAGATTTAATGACATTTCTCTGCGAGTAAGT 526
Qy 160 SerLeuLeuSerSerIleGlnGluHisGluAsnIleIleGlyAspIleSerIleSerLeu 179
Db 527 ACCTTGTTTCTCAGTCCAGGACATGGGAATGCAATAGATGAATCTCCAAAGTCCTTA 586
Qy 180 ValGlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArg 199
Db 587 ATAAGTTTGAATACCAATGCTTGATTTGCGCTCAACATAGAAATCTGAATGGCAAA 646
Qy 200 ValGlnGluAsnAlaPheLysGlnGlnGluMetArgLysLeuGlnGluArgIleTyr 219
Db 647 ATCCAAGAGAAATACCTTCAACAACAAGAGAAATCAGTAAATTAGAGGCGGTGTTTAC 706
Qy 220 AsnAlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyrLeuGlnGluIle 239
Db 707 AATGTATCAGCAGAAATTTATGCTATGAAGAAGAAACAAGTGCATTTTGGAAACAGAAATA 766
Qy 240 LysGlyGluMetLysLeuLeuAsnAsnIleThrAsnAspLeuArgLeuLysAspTrpGlu 259
Db 767 AAAGGAGAAGTGAAGTACTGAATACATCACTAATGATCTCAGACTGAAAGATTGGGAA 826
Qy 260 HisSerGlnThrLeuLysAsnIleThrLeuLeuGlnGly 272
Db 827 CATTTCTCAGACCTTGAGAAATATCACTTTTAATTCAGGT 865

```

RESULT 8

AAF29165
ID AAF29165 standard; DNA; 1347 BP.

AC AAF29165;

XX 05-APR-2001 (first entry)

DT Human DNA encoding type II scavenger receptor.

DE Human; type II scavenger receptor; antibody; macrophage; ds.

XX Homo sapiens.

OS JP2000312595-A.

PN 14-NOV-2000.

PD 16-AUG-1991; 2000JP-00114233.

PF 27-AUG-1990; 90JP-00222398.

PR 16-AUG-1991; 91JP-00229728.

XX (CHUS) CHUGAI PHARM CO LTD.

XX WPI; 2001-161981/17.

DR P-PSDB; AAB49716.

XX New anti-human scavenger receptor antibody useful for the identification

PT and the determination of macrophages.

XX Example 1; Page 7-9; 13pp; Japanese.

XX This invention relates to a poly or monoclonal antibody which targets a
CC human type I or type II scavenger receptor epitope. The anti-scavenger
CC receptor antibody can be used for the identification and determination of
CC macrophages. The present sequence represents DNA encoding the human type
CC II scavenger receptor

XX Sequence 1347 BP; 420 A; 297 C; 298 G; 332 T; 0 U; 0 Other;

Alignment Scores:

| Pred. No.: | 3.05e-85 | Length: | 1347 |
|------------------------|----------|---------------|------|
| Score: | 1042.50 | Matches: | 197 |
| Percent Similarity: | 87.2% | Conservative: | 41 |
| Best Local Similarity: | 72.2% | Mismatches: | 34 |

Query Match: 50.6% Indels: 1
DB: 4 Gaps: 1

US-10-618-570-2 (1-400) x AAF29165 (1-1347)

Qy 1 MetAlaGlnTrpAspAspPheProAspGlnGlnGlnAspThrAspSerCysThrGluSer 20

Db 47 ATGGAGCAGTGGGATCACTTTTCACAATCAACAGGAGGACACTGATAGCTGCTCCGAATCT 106

Qy 21 ValLysPheAspAlaArgSerValThrAlaLeuLeuProProHisProLysAsnGlyPro 40

Db 107 GTGAAATTTGATGCTCGCTCAATGACAGCTTGTCTTCGGAATCTCTAAACACAGCCCT 166

Qy 41 ThrLeuGlnGluArgMetLysSerTyrLysThrAlaLeuIleThrLeuTyrLeuIleVal 60

Db 167 TCCCTTCAGAGAACTGAAAGTCTTCAAGCTGCATGATTGCCCTTTACTCTCTCGTG 226

Qy 61 PheValValLeuValProIleIleGlyIleValAlaAlaGlnLeuLeuLysTrpGluThr 80

Db 227 TTTGCACTTCTCATCCCTCTCATTTGGAATAGTGCAGCTCAACTTCTTGAAGTGGGAAACG 286

Qy 81 LysAsnCysThrValGlySerValAsnAla---AspIleSerProSerProGluGlyLys 99

Db 287 AAGAAATTCCTCAGTTAGTTCCAACTTAATGCAAAATGATAAATCAAAAGTCTCACGGGAAA 346

Qy 100 GlyAsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsn 119

Db 347 GGAATGACAGCGAAGGAAATGAGATTTCAAGAGCTCTTTATGGAAACACATGAGCAAC 406

Qy 120 MetGluSerArgIleGlnTyrLeuSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsn 139

Db 407 ATGGAGAAGAGAAATCCAGCATATTTTAGACATGGAAGCCCAACCTCATGGACACAGCAT 466

Qy 140 PheGlnAsnPheSerIleThrThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsn 159

Db 467 TTCCAAATTTTCAGCATGACAACTGATCAAAAGATTTAATGACATTTCTTCTGCAGTAAGT 526

Qy 160 SerLeuLeuSerSerIleGlnGluHisGluAsnIleIleGlyAspIleSerLysSerLeu 179

Db 527 ACCTTGTTTCTCAGTCCAGGACATGGAATGCAATAGATGAATCTCCAGTCTCTTA 586

Qy 180 ValGlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArg 199

Db 587 ATAAGTTTGAATACCACTTGTGATTTCAGCTCAACATAGAAATCTGAATGGCAAA 646

Qy 200 ValGlnGluAsnAlaPheLysGlnGlnGluMetArgLysLeuGluArgIleTyr 219

Db 647 ATCCACAGAGAATACCTTCAACAACAAGAGGAAATCAGTAAATTAGAGGAGCGTGTTCAC 706

Qy 220 AsnAlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyrLeuGlnGluIle 239

Db 707 ATGTATCAGCAGAAATTTATGGCTATGAAGAAGAAACAAGTGCATTTTGGAAACAGAAATA 766

Qy 240 LysGlyGluMetLysLeuLeuAsnAsnIleThrAsnAspLeuArgLeuLysAspTrpGlu 259

Db 767 AAAGGAGAAGTGAAGTACTGAATAACATCACTAATGATCTCAGACTGAAAGATTGGGAA 826

Qy 260 HisSerGlnThrLeuLysAsnIleThrLeuLeuGlnGly 272

Db 827 CATTTCTCAGACCTTGAGAAATATCACTTTTAATTCAGGT 865

RESULT 9

AAF43260

ID AAT43260 standard; cDNA to mRNA; 1367 BP.

XX AAT43260;

XX 25-FEB-1997 (first entry)

DT Type II macrophage scavenger receptor gene.

XX Type II macrophage scavenger receptor; SRG; kidney; 293 cell;

XX cell substrate; cell attachment; atherosclerosis; ds.

XX OS Homo sapiens.
 XX FH Key Location/Qualifiers
 XX FT CDS 47..1402
 XX FT /*tag= a

XX WO9638725-A1.
 XX PN 05-DEC-1996.
 XX PP 30-MAY-1996; 96WO-US008081.
 XX PR 30-MAY-1995; 95US-00453117.
 XX PA (SMIK) SMITHLINE BEECHAM CORP.
 XX PI Lyako PG, Elshourbagy NAE, Brawner ME;
 XX DR WPI; 1997-034505/03.
 XX DR P-PSDB; AAW08078.
 XX PT New modified human embryonic kidney 293 cells - transfected with a
 XX FT mammalian scavenger receptor gene to enhance ability to attach to a solid
 XX FT support.

XX Example 1; Page 16-18; 32pp; English.
 XX CC Human type I and type II macrophage scavenger receptor genes (AAT43259
 CC and AAT43260, respectively) can be used to transfect human embryonic
 CC kidney 293 cells. The transfected cells show an enhanced ability to
 CC attach to a solid support and are useful for screening cpds. for
 CC biological activity, or for identifying antagonists of the scavenger
 CC receptor gene, e.g. to develop agents for treatment of atherosclerosis

XX SQ Sequence 1367 BP; 427 A; 298 C; 303 G; 339 T; 0 U; 0 Other;

Alignment Scores:
 Pred. NO.: 3,11e-85 Length: 1367
 Score: 1042.50 Matches: 197
 Percent Similarity: 87.2% Conservative: 41
 Best Local Similarity: 72.2% Mismatches: 34
 Query Match: 50.6% Indels: 1
 DB: 2 Gaps: 1

US-10-618-570-2 (1-400) x AAT43260 (1-1367)

Qy 1 MetAlaGlnThrAspAspPheProAspGlnGlnGluAspThrAspSerCysThrGluSer 20
 Db 67 ATGGAGCAGTGGGATCACTTTTCAATCAACAGGAGGACACTGATAGTGTCTCGAATCT 126
 Qy 21 ValLysPheAspAlaArgSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
 Db 127 GTGAATTTGATGCTCGCTCAATGACAGCTTGTCTTCTCGAATCTTAAAGACGCCCT 186
 Qy 41 ThrLeuGlnGluArgMetLysSerTyrLysThrAlaLeuLeuThrLeuTyrLeuLeuVal 60
 Db 187 TCCCTTCAAGAGAAGAACTGAAGTCTTCAAGCTGCACTGATTCCTTACCTCTCTGTG 246
 Qy 61 PheValValLeuValProIleGlyIleValAlaAlaGlnLeuLeuLysTyrPgluThr 80
 Db 247 TTTGAGTTCTATCCCTCTCATTTGAATAGTGGCAGCTCAACTCTCTGAAAGTGGAAACG 306
 Qy 81 LysAsnCysThrValGlySerValAsnAla---AspLysSerProSerProGluGlyLys 99
 Db 307 AAGAAATGCTCAGTTAGTTCAACTCAATGCAATATATACTCAAGTCTCACGGGAAA 366
 Qy 100 GlyAsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsn 119
 Db 367 GGAATGACAGGAGGAGAAATGAGATTTCAAGAGTCTTTATGGACACATGAGCAAC 426
 Qy 120 MetGluSerArgIleGlnTyrLeuSerAspAsnGluAlaLeuLeuAspAlaLysAsn 139

Db 427 ATGGAGAGAGAAATCCAGCATATTTTAGACATGGAAGCCACCTCATGGACACAGGCAT 486
 Qy 140 PheGlnAsnPheSerIleThrThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsn 159
 Db 487 TTCCAAAATTTTTCAGCATGACAACTGATCAAGATTTTAAATGACATCTTCTGCACGTAAGT 546
 Qy 160 SerLeuLeuSerSerIleGlnGluHisGluAsnIleIleGlyAspIleSerLysSerLeu 179
 Db 547 ACCTTGTCTTCTCAGTCCAGGGACATGGGAATGCAATAGATGAATCTCCAAGTCTCTTA 606
 Qy 180 ValGlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArg 199
 Db 607 ATAAGTTTGAATACCACTTCTGATTTGAGCTCAACATAGAAATCTGATGGGCAAA 666
 Qy 200 ValGlnGluAsnAlaPheLysGlnGlnGluMetArgLysLeuGluGluArgIleTyr 219
 Db 667 ATCCAAGAGAAATACCTTCAACACACAGAGAAATCAGTAATTTAGAGGACGCTGTTTAC 726
 Qy 220 AsnAlaSerAlaGluIleLysSerLeuAspGlnLysGlnValTyrLeuGlnGlnGluLe 239
 Db 727 AATGTATCAGCAGAAATATATGCTTATGAAAGAGACACAAAGTGCATTTTGGAAACAGGAATA 786
 Qy 240 LysGlyGluMetLysLeuLeuAsnAsnIleThrAsnAspLeuArgLeuLysAspTyrGlu 259
 Db 787 AAGGAGAGAGTGAAGTACTGAAATACATCACTAATGATCTCAGACTGAAAGATTGGGAA 846
 Qy 260 HisSerGlnThrLeuLysAsnIleThrLeuLeuGlnGly 272
 Db 847 CATTCAGACCTTGAGAAATATCACTTTAATTCAGGT 885

RESULT 10
 ACAS6662
 ID ACAS6662 standard; cDNA; 1367 BP.
 XX AC ACAS6662;
 XX AC ACAS6662;
 DT 06-JUN-2003 (first entry)
 XX DE Human signalling pathway polynucleotide probe SEQ ID NO 1260.
 XX KW Human; probe; ss; array element; Parkinson's disease;
 KW signalling pathway population; cancer; adenocarcinoma; leukaemia;
 KW immunopathy; AIDS; asthma; neuropathy; Alzheimer's disease; microarray.
 XX OS Homo sapiens.
 XX PN US6500938-B1.
 XX PD 31-DEC-2002.
 XX PF 30-JAN-1998; 98US-00016434.
 XX PR 30-JAN-1998; 98US-00016434.
 XX PA (INCY-) INCYTE GENOMICS INC.
 XX PI Au-Young J, Seilhamer JJ;
 XX DR WPI; 2003-352189/33.
 XX PT Combination of polynucleotide probes, useful as array elements in a
 PT microarray for monitoring the expression of a number of target
 PT polynucleotides.
 XX PS Claim 1; SEQ ID NO 1260; 65pp; English.
 XX CC The invention relates to a combination which, comprises a number of
 CC polynucleotide probes comprising a sequence selected from one of the 1490
 CC sequences mentioned in the specification. The combination is useful as an
 CC array element in a microarray for monitoring the expression of a number
 CC of target polynucleotides. The microarray is particularly useful in the
 CC diagnosis and treatment of cancer and immunopathology and neuropathology.
 CC The microarray is useful in diagnostics and treatment regimens, drug

CC and this bin is a subset. The present sequence represents a polymorphic
CC probe of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=0650093891

| | | | |
|----|-----|---|-----|
| Qy | 120 | MetGluSerArgIleGlnTyrLeuSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsn | 139 |
| Db | 427 | ATGGAGAAGAGAAATCCAGCATATTTTATGACATGGAAGCCAACTCTCGGACACAGACGAT | 486 |
| Qy | 140 | PheGlnAsnPheSerIleThrThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsn | 159 |
| Db | 487 | TTCCAAAATTTCAGCATGACCACTGATCAAGATTTTAATGACATTTCTTCGAGCTTAAGT | 546 |
| Qy | 160 | SerLeuLeuSerSerIleGlnGluHisGluAsnIleIleGlyAspIleSerLysSerLeu | 179 |
| Db | 547 | ACCTGTTTTCCTCAGTCAGGAGCATCGGGAAATGCAATAGATGAATCTCCAAGTCCCTTA | 606 |
| Qy | 180 | ValGlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArg | 199 |
| Db | 607 | ATAAGTTTGAAATACCACATTCCTGTGATTTGCGCTCAACATAGAAAATCTGAATGCGAAA | 666 |
| Qy | 200 | ValGlnGluAsnAlaPheLysGlnGlnGluIleuMetArgLysLeuGluGluArgIleTyr | 219 |
| Db | 667 | ATCCAAGAGAAATACCTTCAACAAACAGAGAAAATCAGTAAATTAGAGGAGCGTGTTTAC | 726 |
| Qy | 220 | AsnAlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyrLeuGluGlnGluIle | 239 |

XX WPI; 2001-161981/17.
 DR P-PSDB; AAB49715.
 XX New anti-human scavenger receptor antibody useful for the identification
 PT and the determination of macrophages.
 XX Example 1; Page 5-7; 13pp; Japanese.
 XX This invention relates to a poly or monoclonal antibody which targets a
 CC human type I or type II scavenger receptor epitope. The anti-scavenger
 CC receptor antibody can be used for the identification and determination of
 CC macrophages. The present sequence represents DNA encoding the human type
 CC I scavenger receptor
 XX
 SQ Sequence 2028 BP; 653 A; 369 C; 434 G; 572 T; 0 U; 0 Other;

Alignment Scores:

| | | | |
|------------------------|----------|---------------|------|
| Pred. No.: | 5,14e-85 | Length: | 2028 |
| Score: | 1042.50 | Matches: | 197 |
| Percent Similarity: | 87.2% | Conservative: | 41 |
| Best Local Similarity: | 72.2% | Mismatches: | 34 |
| Query Match: | 50.6% | Indels: | 1 |
| DB: | 4 | Gaps: | 1 |

US-10-618-570-2 (1-400) x AAF29164 (1-2028)

| | | | |
|----|-----|--|-----|
| Qy | 1 | MetAlaGlnTrpAspPheProAspGlnGlnGluAspThrAspSerCysThrGluSer | 20 |
| Db | 47 | ATGGAGCAGTGGGATCACTTTACAAATCAACAGGAGGACACTGATAGCTGCTCGGAATCT | 106 |
| Qy | 21 | ValLysPheAspAlaArgSerValThrAlaLeuLeuProHisProLysAsnGlyPro | 40 |
| Db | 107 | GTGAATTTGATGCTCGCTCAATGACAGCTTTCCTCCGAATCTTAAAAACAGCCCT | 166 |
| Qy | 41 | ThrLeuGlnGluArgMetLysSerTyrLysThrAlaLeuIleThrLeuIleVal | 60 |
| Db | 167 | TCCCTTCAAGAGAACTAGTCCCTCAAGCTGCACGTGATGCCCTTTACTCTCGTG | 226 |
| Qy | 61 | PheValValLeuValProIleIleGlyIleValAlaAlaGlnLeuLeuLysTrpGluThr | 80 |
| Db | 227 | TTTGAGTTCTCATCCCTCTCATTTGGAATAGTGCAGCTCAACTCCTGNAAGTGGGAACG | 286 |
| Qy | 81 | LysAsnCysThrValGlySerValAsnAla---AspIleSerProSerProGluGlyLys | 99 |
| Db | 287 | AAGAAATTCAGTTAGTTCACTAATGCAATGATATTAACCTCAAAGTCTCACGGGAAA | 346 |
| Qy | 100 | GlyAsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsn | 119 |
| Db | 347 | GGAAATGACAGGAGAGGAAATGAGATTTCAAGAGTCTTTATGGAACACATGAGCAAC | 406 |
| Qy | 120 | MetGluSerArgIleGlnTyrLeuSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsn | 139 |
| Db | 407 | ATGGAGAAGAAATCCAGCATATTTAGACATGGAAGCCAACTCATGGACACAGAGCAT | 466 |
| Qy | 140 | PheGlnAsnPheSerIleThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsn | 159 |
| Db | 467 | TTCCAAAATTCAGCATGACCACTCATCAAGATTTAATGACATCTTCTGAGCGTAAGT | 526 |
| Qy | 160 | SerLeuLeuSerSerIleGlnGlnHisGluAsnIleIleGlyAspIleSerLysSerLeu | 179 |
| Db | 527 | ACCTTGTTTCTCAGTCAGGACATGGGAATGCAATGATGAATCTCCAGTCCCTTA | 586 |
| Qy | 180 | ValGlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArg | 199 |
| Db | 587 | ATAAGTTTGAATACCACATGCTTTCATTTGAGCTCAACATAGAAAATCTGAATGGCAA | 646 |
| Qy | 200 | ValGlnGluAsnAlaPheLysGlnGlnGluMetArgLysLeuGluGluArgIleTyr | 219 |
| Db | 647 | ATCCAAGAGAAATACCTTCAACACCAACAGAGGAAATCAGTAAATTAGAGGCGGTTC | 706 |
| Qy | 220 | AsnAlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyrLeuGluGlnIle | 239 |

| | | | |
|----|-----|---|-----|
| Db | 707 | AATGTATCAGCAGAAATTTATGGCTATGAAAGAGAACAAAGTGCATTTTGGAAACAGGAAATA | 766 |
| Qy | 240 | LysGlyGluMetLysLeuLeuAsnAsnIleThrAsnAspLeuArgLeuLysAspTrpGlu | 259 |
| Db | 767 | AAAGGAGAGTGAAGTACTGAAATAACATCACTAATGATCTCAGACTGAAAAGATTGGGAA | 826 |
| Qy | 260 | HisSerGlnThrLeuLysAsnIleThrLeuLeuGlnGly | 272 |
| Db | 827 | CATTCTCAGACCTTGAGAAATATATCACTTTAATTCAAGGT | 865 |

Search completed: February 27, 2006, 09:40:04
 Job time : 659 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 27, 2006, 09:24:14 ; Search time 1029 Seconds
(without alignments)
3214.530 Million cell updates/sec

Title: US-10-618-570-2

Perfect score: 2059

Sequence: 1 MAQWDFPDQEDTSCTES.....DWKATRVGINFTLRTOKE 400

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/abss/ABSSWEB spool/US10618570/runat 27022006 063901 13209/app query fasta_1
-DB=Published Applications NA Main -QFWT=fastap -SUPFIX=rnpsbm -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -HOST=abs04
-USER=US10618570 SCGN 1 1026 @runat 27022006 063901 13209 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSFBLCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA Main:

1: /cgn2_6/ptodata/1/pubna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubna/US08_PUBCOMB.seq.*
3: /cgn2_6/ptodata/1/pubna/US09_PUBCOMB.seq.*
4: /cgn2_6/ptodata/1/pubna/US09B_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubna/US10A_PUBCOMB.seq.*
6: /cgn2_6/ptodata/1/pubna/US10B_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubna/US10C_PUBCOMB.seq.*
8: /cgn2_6/ptodata/1/pubna/US10D_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubna/US10E_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubna/US11_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|--------------------|
| 1 | 2059 | 100.0 | 5177 | 8 | US-10-618-570-1 |
| 2 | 1042.5 | 50.6 | 1077 | 6 | US-10-326-186-5 |
| 3 | 1042.5 | 50.6 | 1367 | 6 | US-10-305-720-1260 |
| 4 | 1042.5 | 50.6 | 2028 | 7 | US-10-426-262-1 |
| 5 | 1042.5 | 50.6 | 2877 | 8 | US-10-741-600-570 |
| 6 | 1042.5 | 50.6 | 3336 | 8 | US-10-741-600-568 |
| 7 | 1042.5 | 50.6 | 3525 | 8 | US-10-741-600-567 |

| | | | | | | |
|----|--------|------|--------|----|---------------------|--------------------|
| 8 | 1042.5 | 50.6 | 3704 | 6 | US-10-240-965-123 | Sequence 123, Appl |
| 9 | 1042.5 | 50.6 | 3719 | 8 | US-10-741-600-569 | Sequence 569, Appl |
| 10 | 1042.5 | 50.6 | 3791 | 8 | US-10-741-600-566 | Sequence 566, Appl |
| 11 | 994.5 | 48.3 | 1330 | 6 | US-10-393-892-41 | Sequence 41, Appl |
| 12 | 994.5 | 48.3 | 1330 | 6 | US-10-394-882-41 | Sequence 41, Appl |
| 13 | 688 | 33.4 | 459 | 7 | US-10-733-031-1 | Sequence 1, Appl |
| 14 | 678 | 32.9 | 486 | 10 | US-11-093-776-6 | Sequence 6, Appl |
| 15 | 678 | 32.9 | 576 | 10 | US-11-093-776-2 | Sequence 2, Appl |
| 16 | 678 | 32.9 | 604 | 8 | US-10-724-527-3 | Sequence 3, Appl |
| 17 | 670 | 32.5 | 540 | 9 | US-10-893-576-57 | Sequence 57, Appl |
| 18 | 670 | 32.5 | 540 | 9 | US-10-893-576-58 | Sequence 58, Appl |
| 19 | 495.5 | 24.1 | 119036 | 8 | US-10-741-600-17721 | Sequence 17721, A |
| 20 | 282 | 13.7 | 201 | 8 | US-10-741-600-13815 | Sequence 13815, A |
| 21 | 282 | 13.7 | 201 | 8 | US-10-741-600-13824 | Sequence 13824, A |
| 22 | 282 | 13.7 | 201 | 8 | US-10-741-600-13833 | Sequence 13833, A |
| 23 | 282 | 13.7 | 201 | 8 | US-10-741-600-13844 | Sequence 13844, A |
| 24 | 282 | 13.7 | 201 | 8 | US-10-741-600-13854 | Sequence 13854, A |
| 25 | 254 | 12.3 | 201 | 8 | US-10-741-600-13814 | Sequence 13814, A |
| 26 | 254 | 12.3 | 201 | 8 | US-10-741-600-13823 | Sequence 13823, A |
| 27 | 254 | 12.3 | 201 | 8 | US-10-741-600-13832 | Sequence 13832, A |
| 28 | 254 | 12.3 | 201 | 8 | US-10-741-600-13842 | Sequence 13842, A |
| 29 | 254 | 12.3 | 201 | 8 | US-10-741-600-13850 | Sequence 13850, A |
| 30 | 204 | 9.9 | 1485 | 7 | US-10-468-335-1 | Sequence 1, Appl |
| 31 | 204 | 9.9 | 1581 | 7 | US-10-138-588-65 | Sequence 65, Appl |
| 32 | 204 | 9.9 | 3644 | 9 | US-10-468-335-3 | Sequence 3, Appl |
| 33 | 201 | 9.8 | 2885 | 9 | US-10-496-905-443 | Sequence 443, Appl |
| 34 | 178 | 8.6 | 201 | 8 | US-10-741-600-48346 | Sequence 48346, A |
| 35 | 171.5 | 8.3 | 3900 | 7 | US-10-311-623-22 | Sequence 22, Appl |
| 36 | 169 | 8.2 | 743 | 6 | US-10-326-186-7 | Sequence 7, Appl |
| 37 | 168.5 | 8.2 | 713 | 6 | US-10-326-186-9 | Sequence 9, Appl |
| 38 | 167 | 8.1 | 1176 | 9 | US-10-671-995A-5 | Sequence 5, Appl |
| 39 | 166.5 | 8.1 | 1395 | 6 | US-10-312-245-5 | Sequence 5, Appl |
| 40 | 155 | 7.5 | 201 | 8 | US-10-741-600-13811 | Sequence 13811, A |
| 41 | 155 | 7.5 | 201 | 8 | US-10-741-600-13820 | Sequence 13820, A |
| 42 | 155 | 7.5 | 201 | 8 | US-10-741-600-13829 | Sequence 13829, A |
| 43 | 155 | 7.5 | 201 | 8 | US-10-741-600-13838 | Sequence 13838, A |
| 44 | 155 | 7.5 | 201 | 8 | US-10-741-600-13849 | Sequence 13849, A |
| 45 | 155 | 7.5 | 840 | 9 | US-10-471-422-3 | Sequence 3, Appl |

ALIGNMENTS

RESULT 1

US-10-618-570-1
; Sequence 1, Application US/10618570
; Publication No. US20040185059A1
; GENERAL INFORMATION:
; APPLICANT: Yla-Herttuala, Seppo
; APPLICANT: Kulomaa, Markku
; APPLICANT: Lehtolainen, Paulina
; APPLICANT: Marjomaki, Varpu
; APPLICANT: Airene, Kari
; TITLE OF INVENTION: Biotin-Binding Receptor Molecules
; FILE REFERENCE: GJE-48
; CURRENT APPLICATION NUMBER: US/10/618,570
; CURRENT FILING DATE: 2003-07-11
; PRIOR APPLICATION NUMBER: US/09/622,804
; PRIOR FILING DATE: 2002-12-10
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 1
; LENGTH: 5177
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Recombinant cDNA encoding fusion protein
; NAME/KEY: CDS
; LOCATION: (1071)..(2270)
; OTHER INFORMATION:
US-10-618-570-1
Alignment Scores:

```
Pred. No.: 1.34e-194 Length: 5177
Score: 2059.00 Matches: 400
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

US-10-618-570-2 (1-400) x US-10-618-570-1 (1-5177)

Qy 1 MetAlaGlnTrpAspPheProAspGlnGlnGluAspThrAspSerCysThrGluSer 20
Db 1071 ATGGCACAGTGGGATGACTTCTCATCAGCAAGAGGACACTGACAGCTGTACAGAGTCT 1130

Qy 21 ValLysPheAspAlaArgSerValThrAlaLeuLeuProHisLeuProLysAsnGlyPro 40
Db 1131 GTGAAGTTCGATCGCTCGCTCAGTGACGCTTGCTTCTCCCATCTCTAAATAATGGCCCA 1190

Qy 41 ThrLeuGlnGluArgMetLysSerTyrLysThrAlaLeuLeuThrLeuTyrLeuLeuVal 60
Db 1191 ACTCTTCAAGAGAGGATGAAGTCTTATAAACTGCACTGATCACCTTTATCTCATTTGTG 1250

Qy 61 PheValValLeuValProIleIleGlyLeuValAlaAlaGlnLeuLeuLysTrpGluThr 80
Db 1251 TTTGTAGTTCGTGCTCCCATCATTTGGCATAGTGGCAGCTCAGCTCTCTGAAATGGGAAACG 1310

Qy 81 LysAsnCysThrValGlySerValAsnAlaAspIleSerProSerProGluGlyLysGly 100
Db 1311 AAGAAATTCACGGTTGGCTCAGTTAATGCAGATATATCTCCAAGTCCGGAAGGCAAGGA 1370

Qy 101 AsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsnMet 120
Db 1371 AATGSCAGTGAAGATGAATGAGATTTTCGAGAGGCTGTGATGGAACGATCAGCAACATG 1430

Qy 121 GluSerArgIleGlnTyrLeuSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsnPhe 140
Db 1431 GAAAGCAGAAATCCAGTATCTTTACAGTAATGAAGCCAATCTCCTAGATGCTTAAGAAATTC 1490

Qy 141 GlnAsnPheSerIleThrThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsnSer 160
Db 1491 CAAAATTTTCAGATAAACAACTGATCAAGATTTAATGATGTTCTTTTCCAGCTAAATTC 1550

Qy 161 LeuLeuSerSerIleGlnGluHisGluAsnIleIleGlyAspIleSerLysSerLeuVal 180
Db 1551 TTACTTTCTCCATCCAGGACATCAGAAATATCATAGGGGATATCTCCAAGTCAATAGTA 1610

Qy 181 GlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArgVal 200
Db 1611 GGTCTGAAACACCAACAGTACTTGAATTCGAGTTTCAGTATTGAAACACTGAATGGCAGAGTC 1670

Qy 201 GlnGluAsnAlaPheLysGlnGlnGluMetArgLysLeuGluGluArgIleTyrAsn 220
Db 1671 CAAAGATGTCATTTAAACAACAGAGAGATGCGTAATTAGAGGCGGTATATACAAAT 1730

Qy 221 AlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyrLeuGluGlnGluIleLys 240
Db 1731 GCATCAGCAGAAATTAAGTCTCTAGATGMAAAACAAGTATATTTGGACAGGAAATATAA 1790

Qy 241 GlyGluMetLysLeuLeuAsnAsnIleThrAsnAspLeuArgLeuLysAspTrpGluHis 260
Db 1791 GGGGAAATCAAACTCTTGAATAATATCAATATGATCTGAGGCTGAAGGATGGGAACAT 1850

Qy 261 SerGlnThrLeuLysAsnIleThrLeuLeuGlnGlyAlaArgLysCysSerLeuThrGly 280
Db 1851 TCTCAGACATTTGAAATAATATACATTTACTCCAAGGTGCCAGAAAGTGTCTGCTGACTGGG 1910

Qy 281 LysTrpThrAsnAspLeuGlySerAsnMetThrIleGlyAlaValAsnSerArgGlyGlu 300
Db 1911 AAATGACCAACGATCTGGGCTCCAAATGACCATCGGGGCTGTGAAACAGCAGAGGTGAA 1970

Qy 301 PheThrGlyThrTyrIleThrAlaValThrAlaThrSerAsnGluIleLysGluSerPro 320
Db 1971 TTCACAGGCACCTACATCACAGCCGTAACAGCCACATCAATGATGATCAAGAGTCAACCA 2030
```

```
Qy 321 LeuHisGlyThrGlnAsnThrIleAsnLysArgThrGlnProThrPheGlyPheThrVal 340
Db 2031 CTGCATGGGACACAAAACACCATCAACAGAGGACCAGCCACCTTTGGCTTCAACCGTC 2090

Qy 341 AsnTrpLysPheSerGluSerThrThrValPheThrGlyGlnCysPheIleAspArgAsn 360
Db 2091 AATTGGAAGTTTTCAGAGTCCACCACCTCTCTTCACGGGCCAGTGTCTTCATAGACAGGAAT 2150

Qy 361 GlyLysGluValLeuLysThrMetTrpLeuLeuArgSerValAsnAspIleGlyAsp 380
Db 2151 GGGAAAGAGGTCCTGAAGACCATGTGCTGCTGGGTCAAGTGTAAATGACATTTGGTAT 2210

Qy 381 AspTrpLysAlaThrArgValGlyIleAsnIlePheThrArgLeuArgThrGlnLysGlu 400
Db 2211 GACTGGAAAGCTACCAGGTCGGCATCAACATCTTCACTCGCCTGCGCACACAGAGGAG 2270

RESULT 2
US-10-326-186-5
; Sequence 5, Application US/10326186
; Publication No. US20030119149A1
; GENERAL INFORMATION:
; APPLICANT: Reddy, Pranhitha
; TITLE OF INVENTION: Trimeric Recombinant Polypeptides
; FILE REFERENCE: 3398-A
; CURRENT APPLICATION NUMBER: US/10/326,186
; CURRENT FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 1077
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1077)
; OTHER INFORMATION:
US-10-326-186-5

Alignment Scores:
Pred. No.: 6.01e-94 Length: 1077
Score: 1042.50 Matches: 197
Percent Similarity: 87.2% Conservative: 41
Best Local Similarity: 72.2% Mismatches: 34
Query Match: 50.6% Indels: 1
DB: 6 Gaps: 1

US-10-618-570-2 (1-400) x US-10-326-186-5 (1-1077)

Qy 1 MetAlaGlnTrpAspPheProAspGlnGlnGluAspThrAspSerCysThrGluSer 20
Db 1 ATGGAGCAGTGGGATCACTTTCAATCAACAGAGGACACTGATAGCTGCTCCGAAATCT 60

Qy 21 ValLysPheAspAlaArgSerValThrAlaLeuLeuProHisLeuProLysAsnGlyPro 40
Db 61 GTGAAATTTGATGCTCGCTCAATCAGACGCTTGTCTCTCGAAATCTCTAAACACAGCCCT 120

Qy 41 ThrLeuGlnGluArgMetLysSerTyrLysThrAlaLeuLeuThrLeuTyrLeuVal 60
Db 121 TCCCTTCAAGAGAAACTGGAAGTCTCTCAAAAGCTCACTGATTCGCCCTTACTCTCTCGTG 180

Qy 61 PheValValLeuValProIleIleGlyLeuValAlaAlaGlnLeuLysTrpGluThr 80
Db 181 TTTGCAGTTCTCATCCCTCTCATTTGGATAGTGGCAGCTCAACTCTCTGAAATGGGAAACG 240

Qy 81 LysAsnCysThrValGlySerValAsnAla---AspIleSerProSerProGluGlyLys 99
Db 241 AAGAATTTGCTCAGTTAGTTCACTCAATGCAAAATGATATACTCAAGTCTCACGGGAAAA 300

Qy 100 GlyAsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsn 119
Db 301 GGAATCAGACGGAAGGAAATGAGATTTCAAGAGTCTTTATGGAAACACATGAGCAAC 360

Qy 120 MetGluSerArgIleGlnTyrLeuSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsn 139
```



```

Db 361 ATGGAGAGAGAGATCAGCATATTTTAGCATGGAAGCAACCTCATGGACACAGAGCAT 420
Qy 140 PheGlnAsnPheSerIleThrThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsn 159
Db 421 TTCCAAAATTTTCAGCATGAGCACTCATCAAGATTTAATGACATTTCTTCGACGCTAAGT 480
Qy 160 SerLeuLeuSerSerIleGlnGluHisGluAsnIleIleGlyAspIleSerIleSerLeu 179
Db 481 ACCTTGTTTCTCAGTCAGGAGACATGGGATGCAATAGATGAATATCCCAAGTCCTTA 540
Qy 180 ValGlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArg 199
Db 541 ATAACTTTGAATACCATCTGCTGATTTGCGCTCAACATAGAAAATCTGAATGGCAAA 600
Qy 200 ValGlnGluAsnAlaPheLysGlnGlnGluMetArgLysLeuGluGluArgIleTyr 219
Db 601 ATCCAAAGAGATACCTTCAACAACAGAGGAAATCAGTAAATTAGAGGAGCGTGTTTAC 660
Qy 220 AsnAlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyrLeuGluGlnIle 239
Db 661 AATGATATCAGCAGAAATTTATGCTATGAAGAAGAACAAAGTGCATTTGGAAACAGGAAATA 720
Qy 240 LysGlyGluMetLysLeuLeuAsnIleThrAsnAspLeuAtcGLeuLysAspTrpGlu 259
Db 721 AAAGAGAGAGTGAAGTACTGAATAAACATCACTAATGATCTCAGACTGAAAGATTGGAA 780
Qy 260 HisSerGlnThrLeuLysAsnIleThrLeuLeuGlnGly 272
Db 781 CATCTCAGACCTTGAGAAATATCATCTTATTCAGGT 819

```

RESULT 3

```

US-10-305-720-1260
; Sequence 1260, Application US/10305720
; Publication No. US20040010136A1
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice K.; Seilhamer, Jeffrey J.
; TITLE OF INVENTION: Composition for the Detection of Signaling Pathway Gene Expression
; FILE REFERENCE: PA-0002-1 CON
; CURRENT APPLICATION NUMBER: US/10/305,720
; CURRENT FILING DATE: 2002-11-26
; PRIOR APPLICATION NUMBER: 09/016,434
; PRIOR FILING DATE: 1998-01-30
; NUMBER OF SEQ ID NOS: 1490
; SOFTWARE: PERL Program
; SEQ ID NO 1260
; LENGTH: 1367
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: GenBank ID No. US20040010136A1 g219991
US-10-305-720-1260

```

```

Alignment Scores:
Pred. No.: 8,466-94 Length: 1367
Score: 1042.50 Matches: 197
Percent Similarity: 87.2% Conservative: 41
Best Local Similarity: 72.2% Mismatches: 34
Query Match: 50.6% Indels: 1
DB: 6 Gaps: 1

```

US-10-618-570-2 (1-400) x US-10-305-720-1260 (1-1367)

```

Qy 1 MetalGlnTrpAspAspPheProAspGlnGlnGluAspThrAspSerCysThrGluSer 20
Db 67 ATGGAGCAGTGGGATCCTTTTCAATCAACAGGAGGACCTGATAGTCTCGAAATCT 126
Qy 21 ValLysPheAspAlaArgSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
Db 127 GTGAATTTGATGCTCGCTCATATGACAGCTTGCTCTCGAATCTCGAATCCATAAAGACGCCCT 186
Qy 41 ThrLeuGlnGluArgMetLysSerTyrLysThrAlaLeuIleThrLeuTyrLeuIleVal 60

```

```

Db 187 TCCCTTCAAGAGAACTGAAGTCTCTCAAGAGCTGCCTGATTCCTTACTCTCTCGTG 246
Qy 61 PheValValLeuValProIleIleGlyIleValAlaalaGlnLeuLeuLysTrpGluThr 80
Db 247 TTTGACGTTTCATCCCTCTCATTTGGAATAGTGGCAGCTCAACTCTCTGAAGTGGGAACG 306
Qy 81 LysAsnCysThrValGlySerValAsnAla---AspIleSerProSerProGluGlyLys 99
Db 307 AAGAATTTGCTCAGTTAGTTCAACTAATGCAATATATATACTCAAGTCTCACGGGAAAA 366
Qy 100 GlyAsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsn 119
Db 367 GGAATGACACGAGAGGAAATGAGATTTCAAGAGTCTTTATGGAAACATAGAGCAAC 426
Qy 120 MetGluSerArgIleGlnTyrLeuSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsn 139
Db 427 ATGGAGAAGAAATCCAGCATATTTTAGACATCGAGGCCAACCTCATGCACACAGAGCAT 486
Qy 140 PheGlnAsnPheSerIleThrThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsn 159
Db 487 TTCCAAAATTTTCAGCATGACAACTGATCAAAAGATTTTATGACATTTCTTCTGCAGCTAAGT 546
Qy 160 SerLeuLeuSerSerIleGlnGluHisGluAsnIleIleGlyAspIleSerIleSerLeu 179
Db 547 ACCTTGTTTTCTCAGTCCAGGAGCATGGGAATGCAATAGATGAATATCTCAAGTCTCTTA 606
Qy 180 ValGlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArg 199
Db 607 ATAAGTTTGAATACCACTGCTTGTATTTGACGCTCAACATAGAAAATCTGAATGGCAAA 666
Qy 200 ValGlnGluAsnAlaPheLysGlnGlnGluMetArgLysLeuGluGluArgIleTyr 219
Db 667 ATCCAAAGAGATACCTTCAACAACAGAGGAAATCAGTAAATTAGAGGAGCGTGTTTAC 726
Qy 220 AsnAlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyrLeuGluGlnGluIle 239
Db 727 AATGATATCAGCAGAAATTTATGCTATGAAGAAGAACAAAGTGCATTTGGAAACAGGAAATA 786
Qy 240 LysGlyGluMetLysLeuLeuAsnIleThrAsnAspLeuArgLeuLysAspTrpGlu 259
Db 787 AAAGAGAGAGTGAAGTACTGATGAATACATCACTAATGATCTCAGACTGAAAGATTGGAA 846
Qy 260 HisSerGlnThrLeuLysAsnIleThrLeuLeuGlnGly 272
Db 847 CATCTCAGACCTTGAGAAATATCATCTTATTCAGGT 885

```

RESULT 4

```

US-10-426-262-1
; Sequence 1, Application US/10426262
; Publication No. US20040018521A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jianfeng
; APPLICANT: Meyers, Deborah
; APPLICANT: Zheng, Sigun
; APPLICANT: Walsh, Patrick C.
; APPLICANT: Isaacs, William B.
; APPLICANT: Bleeker, Eugene
; TITLE OF INVENTION: MUTATIONS IN THE MACROPHAGE SCAVENGER RECEPTOR 1 GENE ALTER RISK
; TITLE OF INVENTION: OF PROSTATE CANCER, ASTHMA, AND CARDIOVASCULAR DISEASE
; FILE REFERENCE: 9151-23
; CURRENT APPLICATION NUMBER: US/10/426,262
; CURRENT FILING DATE: 2003-05-01
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 2028
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (47)..(1402)
US-10-426-262-1

```

```

; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 570
; LENGTH: 2877
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-600-570

Alignment Scores:
Pred. No.: 2,46e-93 Length: 2877
Score: 1042.50 Matches: 197
Percent Similarity: 87.2% Conservative: 41
Best Local Similarity: 72.2% Mismatches: 34
Query Match: 50.6% Indels: 1
DB: Gaps: 1

US-10-618-570-2 (1-400) x US-10-741-600-570 (1-2877)

Qy 1 MetAlaGlnTrpAspAepPheProAspGlnGlnAspThrAspSerCysThrGluSer 20
Db 175 ATGGAGCAGTGGGATCACTTTCAACAATCAACAGGAGGACACTGATAGCTGCTCGGAATCT 234
Qy 21 ValLysPheAspAlaArgSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
Db 235 GTGAAATTTGATGCTCGCTCAATGACAGCTTTGCTTCCTCGGAATCTTAAACACAGCCCT 294
Qy 41 ThrLeuGlnGluArgMetLysSerTyrLysThrAlaLeuLeuThrLeuTyrLeuVal 60
Db 295 TCCCTTCAAGAGAAATGAAAGTCTCTCAAGAGTGCACATGATTCCTTACCTCTCGTG 354
Qy 61 PheValValLeuValProIleGlyLeValAlaAlaGlnLeuLeuLysTrpGluThr 80
Db 355 TTTCAGTTCCTATCCCTCTCATTTGGAATAGTGGCAGCTCAACTCTTGAAGTGGGAAACG 414
Qy 81 LysAsnCysThrValGlySerValAsnAla---AspLysSerProSerProGluGlyLys 99
Db 415 AAGAAATTCGCTCAGTTAGTTCAACTTAATGCAATATGATATAACTCAAGTCTCACCGGAAAA 474
Qy 100 GlyAsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsn 119
Db 475 GGAATGACAGCGAAGAGGAAATGAGATTTCAAGAGCTCTTTATGGAAACACATAGGCAAC 534
Qy 120 MetGluSerArgIleGlnTyrLeuSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsn 139
Db 535 ATGGAGAGAGAAATCCAGCATATTTTAGACATGGAAGCCCACTCATGGACACAGCAT 594
Qy 140 PheGlnAsnPheSerIleThrThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsn 159
Db 595 TTCAAAATTTTCAGCATGACACACTGATCAAAAGATTTAATGACATTTCTTCTGCAGCTAAGT 654
Qy 160 SerLeuLeuSerSerIleGlnGluHisGluAsnIleIleGlyAspIleSerLysSerLeu 179
Db 655 ACCTTGTTCCTCAGTCCAGGAGCATGGGAAATGCAATAGATGAATATCTCAAGTCTCTTA 714
Qy 180 ValGlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArg 199
Db 715 ATAAGTTTGAATACCATTCATTTGATTTCAGCTCAACATAGAAATCTGAATGGCAAA 774
Qy 200 ValGlnGluAsnAlaPheLysGlnGlnGluMetArgLysLeuGluGluArgIleTyr 219
Db 775 ATCCAGAGAAATACCTTCAAAACAAACAGAGGAAATCAGTAAATTTAGAGGAGCGTGTTCAC 834
Qy 220 AsnAlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyrLeuGluGluIle 239
Db 835 AATGTATCRGAGAAATTTATGGCTATGAAAGAGAAACAGTGCAATTTGGACACAGGAATA 894
Qy 240 LysGlyGluMetLysLeuLeuAsnIleThrAsnAspLeuArgLeuLysAspTrpGlu 259

```

RESULT 5

```

US-10-741-600-570
; Sequence 570, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.

```

```
Db 895 AAAGGAGAGTGAAGTACTGAATAACATCACTAATGATCTCAGACTGAAGATTGGAA 954
Qy 260 HisSerGlnThrLeuLysAsnIleThrLeuLeuGlnGly 272
Db 955 CATTCTCAGACCTTGAGAAATATCACTTTAATTCAGGT 993

RESULT 6.
US-10-741-600-568
; Sequence 568, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 568
; LENGTH: 3336
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-600-568

Alignment Scores:
Pred. No.: 3,04e-93 Length: 3336
Score: 1042.50 Matches: 197
Percent Similarity: 87.2% Conservative: 41
Best Local Similarity: 72.2% Mismatches: 34
Query Match: 50.6% Indels: 1
DB: 8 Gaps: 1

US-10-618-570-2 (1-400) x US-10-741-600-568 (1-3336)
Qy 1 MetAlaGlnTrpAspAspPheProAspGlnGlnAspThrAspSerCysThrGluSer 20
Db 160 ATGGAGCAGTGGGATCACTTTCAATCAATCAACAGGAGGACCTGATAGCTGCCGAATCT 219
Qy 21 ValLysPheAspAlaArgSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
Db 220 GTGAATTTGATGCTCGTCAATGACAGCTTGTCTCTCGAATCTTAAACACAGCCT 279
Qy 41 ThrLeuGlnGluArgMetLysSerTyrLysThrAlaLeuIleThrLeuTyrLeuVal 60
Db 280 TCCCTTCAAGAGAACTCAAGTCTTCAAGCTGCACTGATTCCTTACCTCTCTCGTG 339
Qy 61 PheValValLeuValProIleIleGlyIleValAlaAlaGlnLeuLeuTyrTrpGluThr 80
Db 340 TTTCAGTTCATCCCTCTCATTTGGAATAGTGGCAGCTCAACTCTTGAAGTGGGAACG 399
Qy 81 LysAsnCysThrValGlySerValAsnAla---AspIleSerProSerProGluGlyLys 99
Db 400 AAGAAATGCTCAGTTAGTTCACTAATGCAATGATATGATCAAGTCTCACGGGAAA 459
Qy 100 GlyAsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsn 119
Db 460 GGAAATGACGGAAGAGAAATGAGATTTCAAGAGTCTTTATGGAACACATCAGCAAC 519
Qy 120 MetGluSerArgIleGlnTyrLeuSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsn 139
Db 520 ATGGAGAGAGAAATCCAGCATATTTAGCATGGAAGCAACCTCATGGAACACAGAGCAT 579
Qy 140 PheGlnAsnPheSerIleThrThrAspGlnArgPheAsnValLeuPheGlnLeuAsn 159
Db 580 TTCCAAATTTTCAGATCAACACTGATCAAGATTTAATGATCTTTCTCGACCTAAGT 639
Qy 160 SerLeuLeuSerSerIleGlnGlnHisGluAsnIleIleGlyAspIleSerLysSerLeu 179
Db 640 ACCTTGTTTCTCAGTCCAGGACATGGGAATGCAATAGATGAATCTTCAAGTCCCTTA 699
Qy 180 ValGlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArg 199
Db 942 CATTCTCAGACCTTGAGAAATATCACTTTAATTCAGGT 993
```

```
Db 700 ATAAGTTTGAATACCACTGCTTGTGATTGCGCTCAACATAGAAAATCTGAATGGCAA 759
Qy 200 ValGlnGluAsnAlaPheLysGlnGlnGluMetArgLysLeuGluGluArgIleTyr 219
Db 760 ATCCAAGAGAAATACCTTTCAACAAACAAGAGAAATCAGTAATATTAGAGGAGCGTGTTC 819
Qy 220 AsnAlaSerAlaGluIleLysSerLeuAspGlnLysGlnValTyrLeuGluGluIle 239
Db 820 AATGTATCRGAGAAATTTATGGCTATGAAAGAAACAACAGTGCATTTGGAAACAGGAATA 879
Qy 240 LysGlyGluMetLysLeuLeuAsnAsnIleThrAsnAspLeuArgLeuLysAspTyrGlu 259
Db 880 AAAGAGAGTGAAGTACTGATCAATCACTAATGATCTCAGACTGAAAGATTGGAA 939
Qy 260 HisSerGlnThrLeuLysAsnIleThrLeuLeuGlnGly 272
Db 940 CATTCTCAGACCTTGAGAAATATCACTTTAATTCAGGT 978

RESULT 7
US-10-741-600-567
; Sequence 567, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 567
; LENGTH: 3525
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-600-567

Alignment Scores:
Pred. No.: 3,29e-93 Length: 3525
Score: 1042.50 Matches: 197
Percent Similarity: 87.2% Conservative: 41
Best Local Similarity: 72.2% Mismatches: 34
Query Match: 50.6% Indels: 1
DB: 8 Gaps: 1

US-10-618-570-2 (1-400) x US-10-741-600-567 (1-3525)
Qy 1 MetAlaGlnTrpAspAspPheProAspGlnGlnAspThrAspSerCysThrGluSer 20
Db 160 ATGGAGCAGTGGGATCACTTTCAATCAATCAACAGGAGGACCTGATAGCTGCCGAATCT 219
Qy 21 ValLysPheAspAlaArgSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
Db 220 GTGAATTTGATGCTCGTCAATGACAGCTTGTCTCTCGAATCTTAAACACAGCCT 279
Qy 41 ThrLeuGlnGluArgMetLysSerTyrLysThrAlaLeuIleThrLeuTyrLeuVal 60
Db 280 TCCCTTCAAGAGAACTCAAGTCTTCAAGCTGCACTGATTCCTTACCTCTCTCGTG 339
Qy 61 PheValValLeuValProIleIleGlyIleValAlaAlaGlnLeuLeuTyrTrpGluThr 80
Db 340 TTTCAGTTCATCCCTCTCATTTGGAATAGTGGCAGCTCAACTCTTGAAGTGGGAACG 399
Qy 81 LysAsnCysThrValGlySerValAsnAla---AspIleSerProSerProGluGlyLys 99
Db 400 AAGAAATGCTCAGTTAGTTCACTAATGCAATGATATGATCAAGTCTCACGGGAAA 459
Qy 100 GlyAsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsn 119
Db 460 GGAAATGACGGAAGAGAAATGAGATTTCAAGAGTCTTTATGGAACACATCAGCAAC 519
Qy 120 MetGluSerArgIleGlnTyrLeuSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsn 139
Db 942 CATTCTCAGACCTTGAGAAATATCACTTTAATTCAGGT 978
```


ORGANISM: Homo sapiens
US-10-741-600-569

Alignment Scores:
Pred. No.: 3,566-93 Length: 3719
Score: 1042.50 Matches: 197
Percent Similarity: 87.2% Conservatives: 41
Best Local Similarity: 72.2% Mismatches: 34
Query Match: 50.6% Indels: 1
DB: 8 Gaps: 1

US-10-618-570-2 (1-400) x US-10-741-600-569 (1-3719)

```

Qy 1 MetAlaGlnTrpAspAspPheProAspGlnGlnGlnAspThrAspSerCysThrGluSer 20
Db 175 ATGGAGCAGTGGGATCACTTTTCAAAATCAACAGGAGGACACTGATAGCTGCTCGGAATCT 234

Qy 21 VallysPheAspAlaArgSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
Db 235 GTGAATTTGATGCTCGCTCAATGACAGCTTTGCTTCTCCGAATCTCAAAACAGCCCT 294

Qy 41 ThrLeuGlnGluArgMetLysSerTyrLysThrAlaLeuLeuThrLeuTyrLeuVal 60
Db 295 TCCCTTCAAGAGAACTGAGTCTTCAAGCTGACCTGATGCTTACTTCTCTCGTG 354

Qy 61 PheValValLeuValProIleGlyLeValAlaAlaGlnLeuLeuTyrTrpGluThr 80
Db 355 TTTCAGTCTCATCCCTCTCATTTGGAATAGTGGCAGCTCAACTCTCTGAAGTGGGAACG 414

Qy 81 LysAsnCysThrValGlySerValAsnAla---AspIleSerProSerProGluGlyLys 99
Db 415 AAGAAATTCCTCAGTGTAGTTCAACTAATCAATGATATATATCACTCAAGTCTCACCGGAAA 474

Qy 100 GlyAsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsn 119
Db 475 GGAATGACAGGAGAGAAATGAGATTTCAAGAGTCTTTAAGAACACATGAGCAAC 534

Qy 120 MetGluSerArgIleGlnTyrLysSerAspAsnGlnAlaAsnLeuLeuAspAlaLysAsn 139
Db 535 ATGGAGAAGAAATCCAGCATATTTAGACATGGAAGCAACCTCATGGACACAGAGCAT 594

Qy 140 PheGlnAsnPheSerIleThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsn 159
Db 595 TTCCAAAATTTTCAGCATGACAACTGATCAAAAGATTTAATGACATTTCTTCGACGTAAGT 654

Qy 160 SerLeuLeuSerSerIleGlnGluHisGluAsnIleGlyAspIleSerLysSerLeu 179
Db 655 ACCTTGTTTCTCAGTCCAGGACATGGGAATGCAATAGATGAAATCTCCAGTCCCTTA 714

Qy 180 ValGlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArg 199
Db 715 ATAAGTTTGAATACCACATGCTTGTATTTGACGCTCAACATAGAAAATCTCAATGGCAAA 774

Qy 200 ValGlnGluAsnAlaPheLysGlnGlnGluMetArgLysLeuGluGluArgIleTyr 219
Db 775 ATCCAAAGAGATACCTTCAAAACAAAGAGGAAATCAGTAAATTTAGAGGACGCTGTTTAC 834

Qy 220 AsnAlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyrLeuGluGlnGluLe 239
Db 835 AATGTATCRGAGAAATTTATGCTATGAAAGAACAGCAAGTGCAATTTGGAACAGGAATA 894

Qy 240 LysGlyGluMetLysLeuAsnAsnIleThrAsnAspLeuArgLeuLysAspTrpGlu 259
Db 895 AAAGAGAGAGTGAAGTACTGAATAACATCACTAATGATCTCAGACTGAAAGATGGAA 954

Qy 260 HisSerGlnThrLeuLysAsnIleThrLeuLeuGlnGly 272
Db 955 CATCTCAGACCTTGAGAAATATCACTTTAATTCAGGT 993

```

RESULT 10

US-10-741-600-566
; Sequence 566, Application US/10741600
; Publication No. US20050026169A1

GENERAL INFORMATION:

; APPLICANT: CARGILL, Michele et al.
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 ; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001499
 ; CURRENT APPLICATION NUMBER: US/10/741,600
 ; CURRENT FILING DATE: 2003-12-22
 ; NUMBER OF SEQ ID NOS: 73997
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 566
 ; LENGTH: 3791
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-10-741-600-566

Alignment Scores:
 Pred. No.: 3,566-93 Length: 3791
 Score: 1042.50 Matches: 197
 Percent Similarity: 87.2% Conservatives: 41
 Best Local Similarity: 72.2% Mismatches: 34
 Query Match: 50.6% Indels: 1
 DB: 8 Gaps: 1

US-10-618-570-2 (1-400) x US-10-741-600-566 (1-3791)

```

Qy 1 MetAlaGlnTrpAspAspPheProAspGlnGlnGlnAspThrAspSerCysThrGluSer 20
Db 160 ATGGAGCAGTGGGATCACTTTTCAAAATCAACAGGAGGACACTGATAGCTGCTCGGAATCT 219

Qy 21 VallysPheAspAlaArgSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
Db 220 GTGAATTTGATGCTCGCTCAATGACAGCTTTGCTTCTCCGAATCTCAAAACAGCCCT 279

Qy 41 ThrLeuGlnGluArgMetLysSerTyrLysThrAlaLeuLeuThrLeuTyrLeuVal 60
Db 280 TCCCTTCAAGAGAACTGAGATTTCAAGAGTCTTTAAGAACACATGAGCAAC 339

Qy 61 PheValValLeuValProIleGlyLeValAlaAlaGlnLeuLeuTyrTrpGluThr 80
Db 340 TTTCAGTCTCATCCCTCTCATTTGGAATAGTGGCAGCTCAACTCTCTGAAGTGGGAACG 399

Qy 81 LysAsnCysThrValGlySerValAsnAla---AspIleSerProSerProGluGlyLys 99
Db 400 AAGAAATTTTCAGTGTAGTTCAACTAATGCAATGATATATACTCAAGTCTCACCGGAAA 459

Qy 100 GlyAsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsn 119
Db 460 GGAATGACAGGAGAGAAATGAGATTTCAAGAGTCTTTAAGAACACATGAGCAAC 519

Qy 120 MetGluSerArgIleGlnTyrLysSerAspAsnGlnAlaAsnLeuLeuAspAlaLysAsn 139
Db 520 ATGGAGAAGAAATCCAGCATATTTTAGACATGGAAGCAACCTCATGGACACAGAGCAT 579

Qy 140 PheGlnAsnPheSerIleThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsn 159
Db 580 TTCCAAAATTTTCAGCATGACAACTGATCAAAAGATTTAATGACATTTCTTCGACGTAAGT 639

Qy 160 SerLeuLeuSerSerIleGlnGluHisGluAsnIleGlyAspIleSerLysSerLeu 179
Db 640 ACCTTGTTTCTCAGTCCAGGAGACATGGGAATGCAATAGATGAAATCTCCAGTCCCTTA 699

Qy 180 ValGlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArg 199
Db 700 ATAAGTTTGAATACCACATGCTTGTATTTGACGCTCAACATAGAAAATCTGAATGGCAAA 759

Qy 200 ValGlnGluAsnAlaPheLysGlnGlnGluMetArgLysLeuGluGluArgIleTyr 219
Db 760 ATCCAAAGAGATACCTTCAAAACAAAGAGGAAATCAGTAAATTTAGAGGACGCTGTTTAC 819

Qy 220 AsnAlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyrLeuGluGlnGluLe 239
Db 820 AATGTATCRGAGAAATTTATGCTATGAAAGAACAGCAAGTGCAATTTGGAACAGGAATA 879

```

```
Qy 240 LysGlyGluMetLysLeuLeuAsnAsnIleThrAsnAspLeuArgLeuLysAspTrpGlu 259
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 880 AAAGGAGAGTCAAGTACTGATTAACATCACTAATGATCTCAGACTGAAGATTGGAA 939

Qy 260 HisSerGlnThrLysLeuAsnIleThrLeuLeuGlnGly 272
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 940 CATCTCAGACCTTGAGAAATATCACTTTAATTCAAGGT 978

RESULT 11
US-10-393-892-41
; Sequence 41, Application US/10393892
; Publication No. US20030186302A1
; GENERAL INFORMATION:
; APPLICANT: WANG, YIXIN
; TITLE OF INVENTION: COLORECTAL CANCER DIAGNOSTICS
; FILE REFERENCE: CDS 267 US NP
; CURRENT APPLICATION NUMBER: US/10/393,892
; PRIOR FILING DATE: 2003-03-21
; PRIOR APPLICATION NUMBER: 60/368,798
; PRIOR FILING DATE: 2002-03-29
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 41
; LENGTH: 1330
; TYPE: DNA
; ORGANISM: human
US-10-393-892-41

Alignment Scores:
Pred. No.: 5,12e-89 Length: 1330
Score: 994.50 Matches: 189
Percent Similarity: 85.3% Conservative: 44
Best Local Similarity: 69.2% Mismatches: 39
Query Match: 48.3% Indels: 1
DB: 6 Gaps: 1

US-10-618-570-2 (1-400) x US-10-393-892-41 (1-1330)
Qy 1 MetAlaGlnTrpAspPheProAspGlnGlnGluAspThrAspSerCysThrGluSer 20
Db 1 ATGGCGCAGTGGGACAGCTTCACTGATCAACAGGAGGACACTGATAGCTGTTCAGAACTC 60

Qy 21 ValLysPheAspAlaArgSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
Db 61 GTGAAGTTTGATGCTCGCTCCAAATACAGCTTTGCTTCCCAATCTTAAATAATGGCCCT 120

Qy 41 ThrLeuGlnGluArgMetLysSerTyrLysThrAlaLeuIleThrLeuTyrLeuIleVal 60
Db 121 CCACCTTCAAGAAAGCTGAATCTTCAAGCTGCACCTGATTCGCCCTTTATCTCCTTGTG 180

Qy 61 PheValValLeuValProIleIleGlyIleValAlaAlaGlnLeuLeuLysTrpGluThr 80
Db 181 TTTGCTGTCTCATCTTATCATCGCAATATATGGCAGCTCAACTCTCGAAGTGGGAATG 240

Qy 81 LysAsnCysThrValGlySerValAsnAlaAsp---IleSerProSerProGluGlyLys 99
Db 241 AAGAATTGCACAGTTGGTTCAATTAATGCAACAGTGTATCTCTCAGTCTCTCTGGGAAGA 300

Qy 100 GlyAsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsn 119
Db 301 GGAATGACAGTGAAGATGAAGTGAGATTTTCGAGAAGTTGTTATGGAACACATAGCAAG 360

Qy 120 MetGluSerArgIleGlnTyrLeuSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsn 139
Db 361 ATGGAGAAAAGAAATCCAATATATTTTCAGATACTGAAGAAAATCTCGTAGATTCAGAGCAT 420

Qy 140 PheGlnAsnPheSerIleThrThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsn 159
Db 421 TTTCAAAATTTTCAGTGACAACTGATCAACGATTTGCTGATGTCTTCTCCAACTAAGT 480

Qy 160 SerLeuLeuSerSerIleGlnGluHisGluAsnIleIleGlyAspIleSerLysSerLeu 179
Db 481 ACCTTTGTTCCACAGTCCAGGACATGGGAATGCGGTAGATGAATCACCAGGTCTCTTA 540
```

```
Qy 180 ValGlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArg 199
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 541 ATAAGTCTGATATACCACGCTGCTTGATTTGCACCTCTATGTAGAAACACTGAATGTCAA 600

Qy 200 ValGlnGluAsnAlaPheLysGlnGlnGluMetArgLysLeuGluGluArgIleTyr 219
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 601 TTCCAGGAGAAATACACTTAAAGGGCAAGAGAAATCAGCAAAATTAAGGAGCGTGTGCAC 660

Qy 220 AsnAlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyrLeuGluGlnGluIle 239
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 661 AATGCATCAGCAGAAATTTATGTTCTATGAAAGAGAACAAAGTGCACTTTGGAAACAGGAATA 720

Qy 240 LysGlyGluMetLysLeuLeuAsnAsnIleThrAsnAspLeuArgLeuLysAspTrpGlu 259
Db 721 AAAGAGAGTGAAGTCTCGAATTAACATCACTAATGATCTCAGGCTGAAAGATTGGGA 780

Qy 260 HisSerGlnThrLeuLysAsnIleThrLeuLeuGlnGly 272
Db 781 CATCTCAGACCTTGAGAAATATCACTTTAATTCAAGGT 819
```

RESULT 12

```
US-10-394-382-41
; Sequence 41, Application US/10394382
; Publication No. US20030186303A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Wang, Yixin
; TITLE OF INVENTION: COLORECTAL CANCER DIAGNOSTICS
; FILE REFERENCE: CDS 266 US NP
; CURRENT APPLICATION NUMBER: US/10/394,382
; CURRENT FILING DATE: 2003-03-21
; PRIOR APPLICATION NUMBER: 60/368,687
; PRIOR FILING DATE: 2002-03-29
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 41
; LENGTH: 1330
; TYPE: DNA
; ORGANISM: human
US-10-394-382-41
```

```
Alignment Scores:
Pred. No.: 5,12e-89 Length: 1330
Score: 994.50 Matches: 189
Percent Similarity: 85.3% Conservative: 44
Best Local Similarity: 69.2% Mismatches: 39
Query Match: 48.3% Indels: 1
DB: 6 Gaps: 1
```

US-10-618-570-2 (1-400) x US-10-394-382-41 (1-1330)

```
Qy 1 MetAlaGlnTrpAspPheProAspGlnGlnGluAspThrAspSerCysThrGluSer 20
Db 1 ATGGCGCAGTGGGACAGCTTCACTGATCAACAGGAGGACACTGATAGCTGTTCAGAACTC 60

Qy 21 ValLysPheAspAlaArgSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
Db 61 GTGAAGTTTGATGCTCGCTCCAATACAGCTTTGCTTCCCAATCTTAAATAATGGCCCT 120

Qy 41 ThrLeuGlnGluArgMetLysSerTyrLysThrAlaLeuIleThrLeuTyrLeuIleVal 60
Db 121 CCACCTTCAAGAAAGCTGAATCTTCAAGCTGCACCTGATTCGCCCTTTATCTCCTTGTG 180

Qy 61 PheValValLeuValProIleIleGlyIleValAlaAlaGlnLeuLeuLysTrpGluThr 80
Db 181 TTTGCTGTCTCATCTTATCATCGCAATATATGGCAGCTCAACTCTCGAAGTGGGAATG 240

Qy 81 LysAsnCysThrValGlySerValAsnAlaAsp---IleSerProSerProGluGlyLys 99
Db 241 AAGAATTGCAGTGTGGTTCAATTAATGCAACAGTGTATCTCTCAGTCTCTCTGGGAAGA 300

Qy 100 GlyAsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsn 119
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```


Db 301 GGAAATGACAGTGAAGATGAGATTTCGAGAGTTGTTATCGAACACATTAGCAAG 360
Qy 120 MetGluSerArgIleGlnTyrLeuSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsn 139
Db 361 ATGGAGAAAGAAATCCATATATATTTTCAGATCTAGCAAGAAATCTCGTAGATTCAGAGCAT 420
Qy 140 PheGlnAsnPheSerIleThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsn 159
Db 421 TTTCAAATTTTCAGTGTGACAACTGATCAAGATTTGCTGATGTTCTTCTCAACTAAGT 480
Qy 160 SerLeuLeuSerIleGlnGluHisGluAsnIleGlyAspIleSerLysSerLeu 179
Db 481 ACCTTGTTCACAGTCCAGGAGCATGGGAATGCGTAGATGAATCAACAGCTCTT 540
Qy 180 ValGlyLeuAsnThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArg 199
Db 541 ATAAGTCTGAATACACGCTCTTGAATTTGACCTCTATGTAGAAACACTGAATGTCAA 600
Qy 200 ValGlnGluAsnAlaPheLysGlnGluGluMetArgLysLeuGluGluArgIleTyr 219
Db 601 TTCAGAGAGAAATACACTTAAGGGCAAGAGAAATACAGAAATTAAGAGGCGTGTGCAC 660
Qy 220 AsnAlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyrLeuGluGlnGluLe 239
Db 661 AATGCATCAGCAGAAATATGCTCTATGAAAGAAACAGAGTGCATTTGGAAACAGAAATA 720
Qy 240 LysGlyGluMetLysLeuLeuAsnAsnIleThrAsnAspLeuArgLeuLysAspTrpGlu 259
Db 721 AAAAGAGAGTGAAGTCTCTGAATACATCACTAATGATCTCAGGCTGAAAGATTGGAA 780
Qy 260 HisSerGlnThrLeuLysAsnIleThrLeuLeuGlnGly 272
Db 781 CATCTCAGAGTGTGAGAAATATCACTTTAATCAAGT 819

RESULT 13

US-10-733-031-1

; Sequence 1, Application US/10733031
; Publication No. US20040175441A1
; GENERAL INFORMATION:
; APPLICANT: BOOTLAND, LINDA M.
; APPLICANT: BEIFUSS, KATHERINE
; APPLICANT: JILKA, JOSEPH
; APPLICANT: LIZAMA, MALCOLM L.
; APPLICANT: LAMPHEAR, BARRY
; APPLICANT: STREATHFIELD, STEPHEN
; TITLE OF INVENTION: IMMUNIZATION OF FISH WITH PLANT-EXPRESSED RECOMBINANT
; TITLE OF INVENTION: PROTEINS
; FILE REFERENCE: 10044
; CURRENT APPLICATION NUMBER: US/10733,031
; CURRENT FILING DATE: 2003-12-11
; PRIOR APPLICATION NUMBER: 60/433,381
; PRIOR FILING DATE: 2002-12-13
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 1
; LENGTH: 459
; TYPE: DNA
; ORGANISM: Hordeum vulgare
US-10-733-031-1

Alignment Scores:

| | | | |
|------------------------|----------|---------------|-----|
| Pred. No.: | 4.91e-59 | Length: | 459 |
| Score: | 688.00 | Matches: | 130 |
| Percent Similarity: | 98.5% | Conservative: | 1 |
| Best Local Similarity: | 97.7% | Mismatches: | 2 |
| Query Match: | 33.4% | Indels: | 0 |
| DB: | 7 | Gaps: | 0 |

US-10-618-570-2 (1-400) x US-10-733-031-1 (1-459)

Qy 268 ThrLeuLeuGlnGlyAlaArgLysSerLeuThrGlyLysThrAsnAspLeuGly 287
Db 58 TCCCTCCGACGCGCGGAGGAGTGTCTCCCTCACCGCAAGTGGACCAATGACCTCGGC 117

Qy 288 SerAsnMetThrIleGlyAlaValAsnSerArgGlyGluPheThrGlyThrTyrIleThr 307
Db 118 TCCAACATGACCATCGGCGCGTGAACCTCCAGGGCGAGTTCCACGGCACCTACATCAC 177
Qy 308 AlaValThrAlaThrSerAsnGluIleLysGluSerProLeuHisGlyThrGlnAsnThr 327
Db 178 GCGGTGACCGCCACTCAAGAGATCAAGAGTCCCTCCACGGTACCCAGAACACC 237
Qy 328 IleAsnLysArgThrGlnProThrPheGlyPheThrValAsnTrpLysPheSerGluSer 347
Db 238 ATCAACAAGAGGAGCCACCCAGCCACTTCGGCTTCCACCGTGAACCTGGAAGTTCTCCGAGTCC 297
Qy 348 ThrThrValPheThrGlyGlnCysPheIleAspArgAsnGlyLysGluValLeuLysThr 367
Db 298 ACCACCGTGTTCACCGGCCAGTGTCTCATCGACCGCAACGCGCAAGGAGGTCTCAAGACC 357
Qy 368 MetTrpLeuLeuArgSerSerValAsnAspIleGlyAspAspTyrLysAlaThrArgVal 387
Db 358 ATGTGGCTCTCGAGGAGCTCCGTGAATGACATCGCGCAGCATCGGAAGGCCACCGCGTG 417
Qy 388 GlyIleAsnIlePheThrArgLeuArgThrGlnLysGlu 400
Db 418 GGCATCAACATCTTCACCGCCCTCCGACCCAGCAAGGAG 456

RESULT 14

US-11-093-776-6

; Sequence 6, Application US/11093776
; Publication No. US20050172356A1
; GENERAL INFORMATION:
; APPLICANT: Christeller, John Tane
; APPLICANT: Sutherland, Paul William
; APPLICANT: Murray, Colleen
; APPLICANT: Markwick, Ngaire Patricia
; APPLICANT: Phillip, Bruce Allan
; APPLICANT: Malone, Louise Anne
; APPLICANT: Burgess, Elisabeth Phyllis
; APPLICANT: Phung, Margaret Mary
; APPLICANT: Phung, Thai Hong
; APPLICANT: The Horticulture and Food Research Institute of
; TITLE OF INVENTION: Chimeric Polypeptides Allowing Expression of
; TITLE OF INVENTION: Plant-Noxious Proteins
; FILE REFERENCE: 020829-00010005
; CURRENT APPLICATION NUMBER: US/11/093,776
; CURRENT FILING DATE: 2005-03-29
; PRIOR APPLICATION NUMBER: US/09/743,690
; PRIOR FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: NZ 331002
; PRIOR FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: WO PCT/NZ99/00110
; PRIOR FILING DATE: 1999-07-15
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 486
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: PPI-I/Avidin
; OTHER INFORMATION: gene fusion
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(486)
; OTHER INFORMATION: PPI-I/Avidin fusion protein
US-11-093-776-6

Alignment Scores:

| | | | |
|------------------------|----------|---------------|-----|
| Pred. No.: | 5.33e-58 | Length: | 486 |
| Score: | 678.00 | Matches: | 128 |
| Percent Similarity: | 100.0% | Conservative: | 0 |
| Best Local Similarity: | 100.0% | Mismatches: | 0 |
| Query Match: | 32.9% | Indels: | 0 |

```
DB: 10 Gaps: 0
US-10-618-570-2 (1-400) x US-11-093-776-6 (1-486)
Qy 273 AlaArgLysCysSerLeuThrGlyLysTrpThrAsnAspLeuGlySerAsnMetThrIle 292
Db 100 GCCAGAAAGTCTCGCTGACTGGGAAATGGACCAACGATCTGGGCTCCAAACATGACCATC 159
Qy 293 GlyAlaValAsnSerArgGlyGluPheThrGlyThrTyrlleThrAlaValThrAlaThr 312
Db 160 GGGGCTGTGAACAGCAGAGGTGAATTCACAGGCACCTACATCACAGCCGTAAACAGCCACA 219
Qy 313 SerAsnGluIleLysGluSerProLeuHisGlyThrGlnAsnThrIleAsnLysArgThr 332
Db 220 TCNAATGAGATCAAGAGTCAACATTCATGGGACACAAACACATCAACAGAGACC 279
Qy 333 GlnProThrPheGlyPheThrValAsnTrpLysPheSerGluSerThrThrValPheThr 352
Db 280 CAGCCACCTTTGGCTTCACCGTCAATTTGGAAGTTTTCAGAGTCCACCATGTCTTCACG 339
Qy 353 GlyGlnCysPheIleAspArgAsnGlyLysGluValleuLysThrMetTrpLeuLeuArg 372
Db 340 GGCCAGTCTTCATAGACAGAAATGGGAAGAGGTCTGAAAGACCATGTGGCTCTCGG 399
Qy 373 SerSerValAsnAspIleGlyAspAspTrpLysAlaThrArgValGlyIleAsnIlePhe 392
Db 400 TCNAAGTGTATGACATTTGGTGTGATGACTGGAAAGCTACCAAGGTCGGCATCATCATCTTC 459
Qy 393 ThrArgLeuArgThrGlnLysGlu 400
Db 460 ACTCGCTCGGCACACAGAGAGAG 483
```

RESULT 15

```
US-11-093-776-2
; Sequence 2, Application US/11093776
; Publication No. US20050172356A1
; GENERAL INFORMATION:
; APPLICANT: Christeller, John Tane
; APPLICANT: Sutherland, Paul William
; APPLICANT: Murray, Colleen
; APPLICANT: Markwick, Ngaire Patricia
; APPLICANT: Philip, Bruce Allan
; APPLICANT: Malone, Louise Anne
; APPLICANT: Burgess, Elisabeth Phyllis
; APPLICANT: Phung, Margaret Mary
; APPLICANT: The Horticulture and Food Research Institute of
; APPLICANT: New Zealand Limited
; TITLE OF INVENTION: Chimeric Polypeptides Allowing Expression of
; FILE REFERENCE: 020829-000100US
; CURRENT APPLICATION NUMBER: US/11/093,776
; PRIOR FILING DATE: 2005-03-29
; PRIOR FILING DATE: 2001-01-12
; PRIOR FILING DATE: 2001-01-12
; PRIOR FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: WO PCT/NZ99/00110
; PRIOR FILING DATE: 1999-07-15
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 576
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:avidin cdna
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (44)..(502)
; OTHER INFORMATION: avidin (pGEMav)
; FEATURE:
```

```
; NAME/KEY: sig peptide
; LOCATION: (44)..(115)
; OTHER INFORMATION: signal sequence
US-11-093-776-2
```

```
Alignment Scores:
Pred. No.: 6.8e-58 Length: 576
Score: 678.00 Matches: 128
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 32.9% Indels: 0
DB: 10 Gaps: 0
```

US-10-618-570-2 (1-400) x US-11-093-776-2 (1-576)

```
Qy 273 AlaArgLysCysSerLeuThrGlyLysTrpThrAsnAspLeuGlySerAsnMetThrIle 292
Db 116 GCCAGAAAGTCTCGCTGACTGGGAAATGGACCAACGATCTGGGCTCCAAACATGACCATC 175
Qy 293 GlyAlaValAsnSerArgGlyGluPheThrGlyThrTyrlleThrAlaValThrAlaThr 312
Db 176 GGGGCTGTGAACAGCAGAGGTGAATTCACAGGCACCTACATCACAGCCGTAAACAGCCACA 235
Qy 313 SerAsnGluIleLysGluSerProLeuHisGlyThrGlnAsnThrIleAsnLysArgThr 332
Db 236 TCNAATGAGATCAAGAGTCAACATTCATGGGACACAAACACATCAACAGAGACC 295
Qy 333 GlnProThrPheGlyPheThrValAsnTrpLysPheSerGluSerThrThrValPheThr 352
Db 296 CAGCCACCTTTGGCTTCACCGTCAATTTGGAAGTTTTCAGAGTCCACCATGTCTTCACG 355
Qy 353 GlyGlnCysPheIleAspArgAsnGlyLysGluValleuLysThrMetTrpLeuLeuArg 372
Db 356 GGCCAGTCTTCATAGACAGAAATGGGAAGAGGTCTCTGAAAGACCATGTGGCTCTCGG 415
Qy 373 SerSerValAsnAspIleGlyAspAspTrpLysAlaThrArgValGlyIleAsnIlePhe 392
Db 416 TCNAAGTGTATGACATTTGGTGTGATGACTGGAAAGCTACCAAGGTCGGCATCATCATCTTC 475
Qy 393 ThrArgLeuArgThrGlnLysGlu 400
Db 476 ACTCGCTCGGCACACAGAGAGAG 499
```

Search completed: February 27, 2006, 10:42:01
Job time : 1037 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 27, 2006, 09:29:22 ; Search time 507 Seconds

(without alignments)
1682.341 Million cell updates/sec

Title: US-10-618-570-2

Perfect score: 2059

Sequence: 1 MAQWDDPPDQEDTDSCTES.....DWKATRVGINFTRLRTQKE 400

Scoring table:

| | BLOSUM62 |
|---------------------------|----------|
| Xgapop 10.0 , Xgapext 0.5 | |
| Ygapop 10.0 , Ygapext 0.5 | |
| Fgapop 6.0 , Fgapext 7.0 | |
| Delop 6.0 , Delext 7.0 | |

Searched: 7209121 seqs, 1066183437 residues

Total number of hits satisfying chosen parameters: 14418242

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/abss/ABSSWEB.spool/US10618570/runat 27022006 063904 13261/app query.fasta_1
-DB=Published Applications NA New -QFMT=fastap -SUFFIX=rnpbn -MINMATCH=0.1
-LOOCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=ptc -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -HOST=abs05p
-USER=US10618570 @CGN 1 1 335 @runat 27022006 063904 13261 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSFBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THRRADS=1 -XGAPOPT=0.5 -FGAPOPT=6 -FGAPEXT=7
-YGAPOPT=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA New:

| | 1: | 2: | 3: | 4: | 5: | 6: | 7: | 8: | 9: | 10: | 11: | 12: | 13: |
|--|----|----|----|----|----|----|----|----|----|-----|-----|-----|-----|
| /cg2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.* | | | | | | | | | | | | | |
| /cg2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.* | | | | | | | | | | | | | |
| /cg2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.* | | | | | | | | | | | | | |
| /cg2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.* | | | | | | | | | | | | | |
| /cg2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.* | | | | | | | | | | | | | |
| /cg2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.* | | | | | | | | | | | | | |
| /cg2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.* | | | | | | | | | | | | | |
| /cg2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.* | | | | | | | | | | | | | |
| /cg2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.* | | | | | | | | | | | | | |
| /cg2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.* | | | | | | | | | | | | | |
| /cg2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.* | | | | | | | | | | | | | |
| /cg2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.* | | | | | | | | | | | | | |
| /cg2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.* | | | | | | | | | | | | | |

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length DB | ID | Description |
|------------|--------|-------------|-----------|----|-------------------------------------|
| 1 | 1016.5 | 49.4 | 2877 | 8 | US-10-995-561-372 Sequence 372, App |
| 2 | 1016.5 | 49.4 | 3336 | 8 | US-10-995-561-370 Sequence 370, App |
| 3 | 1016.5 | 49.4 | 3525 | 8 | US-10-995-561-369 Sequence 369, App |
| 4 | 1016.5 | 49.4 | 3719 | 8 | US-10-995-561-371 Sequence 371, App |

| | 5 | 1016.5 | 49.4 | 3791 | 8 | US-10-995-561-368 | Sequence 368, App |
|--|----|--------|------|--------|----|---------------------|-------------------|
| | 6 | 490.5 | 23.8 | 119036 | 8 | US-10-995-561-13314 | Sequence 13314, A |
| | 7 | 282 | 13.7 | 201 | 8 | US-10-995-561-9974 | Sequence 9974, Ap |
| | 8 | 282 | 13.7 | 201 | 8 | US-10-995-561-10011 | Sequence 10011, A |
| | 9 | 282 | 13.7 | 201 | 8 | US-10-995-561-10046 | Sequence 10046, A |
| | 10 | 282 | 13.7 | 201 | 8 | US-10-995-561-10086 | Sequence 10086, A |
| | 11 | 282 | 13.7 | 201 | 8 | US-10-995-561-10109 | Sequence 10109, A |
| | 12 | 273 | 13.3 | 201 | 8 | US-10-995-561-9940 | Sequence 9940, Ap |
| | 13 | 273 | 13.3 | 201 | 8 | US-10-995-561-9977 | Sequence 9977, Ap |
| | 14 | 273 | 13.3 | 201 | 8 | US-10-995-561-10014 | Sequence 10014, A |
| | 15 | 273 | 13.3 | 201 | 8 | US-10-995-561-10049 | Sequence 10049, A |
| | 16 | 273 | 13.3 | 201 | 8 | US-10-995-561-10100 | Sequence 10100, A |
| | 17 | 272 | 13.2 | 201 | 8 | US-10-995-561-9941 | Sequence 9941, Ap |
| | 18 | 272 | 13.2 | 201 | 8 | US-10-995-561-9978 | Sequence 9978, Ap |
| | 19 | 272 | 13.2 | 201 | 8 | US-10-995-561-10015 | Sequence 10015, A |
| | 20 | 272 | 13.2 | 201 | 8 | US-10-995-561-10051 | Sequence 10051, A |
| | 21 | 272 | 13.2 | 201 | 8 | US-10-995-561-10101 | Sequence 10101, A |
| | 22 | 263 | 12.8 | 201 | 8 | US-10-995-561-9939 | Sequence 9939, Ap |
| | 23 | 263 | 12.8 | 201 | 8 | US-10-995-561-9976 | Sequence 9976, Ap |
| | 24 | 263 | 12.8 | 201 | 8 | US-10-995-561-10013 | Sequence 10013, A |
| | 25 | 263 | 12.8 | 201 | 8 | US-10-995-561-10048 | Sequence 10048, A |
| | 26 | 263 | 12.8 | 201 | 8 | US-10-995-561-10097 | Sequence 10097, A |
| | 27 | 254 | 12.3 | 201 | 8 | US-10-995-561-9973 | Sequence 9973, Ap |
| | 28 | 254 | 12.3 | 201 | 8 | US-10-995-561-10010 | Sequence 10010, A |
| | 29 | 254 | 12.3 | 201 | 8 | US-10-995-561-10045 | Sequence 10045, A |
| | 30 | 254 | 12.3 | 201 | 8 | US-10-995-561-10085 | Sequence 10085, A |
| | 31 | 254 | 12.3 | 201 | 8 | US-10-995-561-10108 | Sequence 10108, A |
| | 32 | 241 | 11.7 | 201 | 8 | US-10-995-561-9956 | Sequence 9956, Ap |
| | 33 | 241 | 11.7 | 201 | 8 | US-10-995-561-9993 | Sequence 9993, Ap |
| | 34 | 241 | 11.7 | 201 | 8 | US-10-995-561-10029 | Sequence 10029, A |
| | 35 | 241 | 11.7 | 201 | 8 | US-10-995-561-10068 | Sequence 10068, A |
| | 36 | 241 | 11.7 | 201 | 8 | US-10-995-561-10092 | Sequence 10092, A |
| | 37 | 241 | 11.7 | 201 | 8 | US-10-995-561-52999 | Sequence 52999, A |
| | 38 | 233 | 11.3 | 286 | 12 | US-11-128-061-1104 | Sequence 1104, Ap |
| | 39 | 233 | 11.3 | 286 | 12 | US-11-128-061-4746 | Sequence 4746, Ap |
| | 40 | 233 | 11.3 | 286 | 12 | US-11-128-049-1104 | Sequence 1104, Ap |
| | 41 | 233 | 11.3 | 286 | 12 | US-11-128-049-4746 | Sequence 4746, Ap |
| | 42 | 220.5 | 10.7 | 201 | 8 | US-10-995-561-9947 | Sequence 9947, Ap |
| | 43 | 220.5 | 10.7 | 201 | 8 | US-10-995-561-9985 | Sequence 9985, Ap |
| | 44 | 220.5 | 10.7 | 201 | 8 | US-10-995-561-10020 | Sequence 10020, A |
| | 45 | 220.5 | 10.7 | 201 | 8 | US-10-995-561-10059 | Sequence 10059, A |

ALIGNMENTS

RESULT 1

US-10-995-561-372
; Sequence 372, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01559
; CURRENT APPLICATION NUMBER: US/10/995.561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 372
; LENGTH: 2877
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-372

| Alignment Scores: | 2.32e-101 | Length: | 2877 |
|---------------------|-----------|---------------|------|
| Pred. No.: | 1016.50 | Matches: | 193 |
| Score: | 85.3% | Conservative: | 40 |
| Percent Similarity: | 70.7% | Mismatches: | 39 |
| Query Match: | 49.4% | Indels: | 1 |
| DB: | 8 | Gaps: | 1 |

```
US-10-618-570-2 (1-400) x US-10-995-561-372 (1-2877)
QY 1 MetAlaGlnTrpAspPheProAspGlnGlnAspThrAspSerCysThrGluSer 20
Db 175 ATGGAGCAGTGGGATCACTTTACAAATCAACAGGAGGACACTGATAGCTGCTCCGAATCT 234
QY 21 ValLysPheAspAlaArgSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
Db 235 GTGAATAATTGATCTCGCTCAATGACAGCTTTGCTTCTCCGAATCTTAAAAAACAGCCCT 294
QY 41 ThrLeuGlnGluArgMetLysSerTyrLysThrAlaLeuLeuLeuLeuLeuVal 60
Db 295 TMCCTTCAAGAGAACTCAAGTCTTCAAGAGTCACTGCTGCTTACCTTCTCCTCGTG 354
QY 61 PheValValLeuValProIleGlyLeuValAlaAlaGlnLeuLeuLysTrpGluThr 80
Db 355 TTTGAGATTCTCATCCCTCTCATTTGGAATAGTGGCAGCTCAACTCTTGAAGTGGGAACG 414
QY 81 LysAsnCysThrValGlySerValAsnAla---AspLysSerProSerProGluGlyLys 99
Db 415 AAGAAATGCTCAGTTAGTTCAACTAATGCAATGATAATACTCAAGTCTCACGGGAAA 474
QY 100 GlyAsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsn 119
Db 475 GGAATGACAGCAAGAGGAAATGAGATTCAAGAAAGCTTTATGGAACACATGAGCAAC 534
QY 120 MetGluSerArgLysGlnTyrLeuSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsn 139
Db 535 ATGGAGAGAGAAATCCAGCATATTGACATATGGAAGCCCACTCATGGACACAGAGCAT 594
QY 140 PheGlnAsnPheSerIleThrThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsn 159
Db 595 TTCCAAAATTTCCAGCATGACAACTGATCAAGATTATATGACATCTTCTGACGTAAGT 654
QY 160 SerLeuLeuSerSerIleGlnGluHisGluAsnIleIleGlyAspLysSerLysSerLeu 179
Db 655 ACCTTGTTTCTCAGTCAGGCAGACATGGGAATGCAATAKATGAAATCTCCAAGTCCTTA 714
QY 180 ValGlyLeuAsnThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArg 199
Db 715 ATAAGTTGAATACCACTATGCTGATTTGACAGTCAACATAGAAAATCTCAATGGCAAA 774
QY 200 ValGlnGluAsnAlaPheLysGlnGlnGluMetArgLysLeuGluGluArgIleTyr 219
Db 775 ATCCAGAGAAATACCTTCAACACACAGGAAATCAGTAATTAGAGGACGCTGTTAC 834
QY 220 AsnAlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyrLeuGluGlnGluIle 239
Db 835 AATGTATCRGAGAAATTTATGCTATGAAAGAAACAAAGTGCATTTTGGAAACAGGAAATA 894
QY 240 LysGlyGluMetLysLeuLeuAsnAsnIleThrAsnAspLeuArgLeuLysAspTrpGlu 259
Db 895 AAAGGAGAAGTGAAGTAGTCAATTAACATCACTAATGATCTCAGACTGAAAGATTGGGAA 954
QY 260 HisSerGlnThrLeuLysAsnIleThrLeuLeuGlnGly 272
Db 955 CATTTCTCAGACCTTGAGAAATATCAATTTAATCAAGT 993

RESULT 2
US-10-995-561-370
; Sequence 370, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 370
```

```

; Sequence 369, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 369
; LENGTH: 3525
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-369

Alignment Scores:
Pred. No.: 3,03e-101 Length: 3525
Score: 1016.50 Matches: 193
Percent Similarity: 85.3% Conservative: 40
Best Local Similarity: 70.7% Mismatches: 39
Query Match: 49.4% Indels: 1
DB: 8 Gaps: 1

US-10-618-570-2 (1-400) x US-10-995-561-369 (1-3525)

Qy 1 MetAlaGlnTTPAspAAspPheProAspGlnGlnAspThrAspSerCysThrGluSer 20
Db 160 ATGAGCAGTGGGATCACTTTTCACATCAACAGAGGACACTGATAGCTGCTCGAATCT 219
Qy 21 ValLysPheAspAlaArgSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
Db 220 GTGAAATTTGATGCTCGCTCAATGACAGCTTTGCTTCCTCGAATCTCAAAACAGCCCT 279
Qy 41 ThrLeuGlnGluArgMetLysSerTyrLysThrAlaLeuLeuThrLeuTyrLeuVal 60
Db 280 TMCCTTCAAGAGAARCTCAAGTCTTCAAAAGCTGCACCTGRTTGCCTTTACCTCTCGTG 339
Qy 61 PheValValLeuValProIleGlyLeuValAlaAlaGlnLeuLeuLysThrGluThr 80
Db 340 TTTGAGTTCTCACTCCCTCTCATTTGGAATAGTGGAGCTCAACTCTCGAATGGGAACG 399
Qy 81 LysAsnCysThrValGlySerValAsnAla---AspIleSerProSerProGluGlyLys 99
Db 400 AAGATTGCTCAGTTAGTTCAACTAATCAATATGATATATACTCAAGTCTCAACGGAAA 459
Qy 100 GlyAsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsn 119
Db 460 GGAATGACACGAAAGAGAAATGAGATTTCAAGAAAGCTTTTATGGAACACATGAGCAAC 519
Qy 120 MetGluSerArgIleGlnTyrLeuSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsn 139
Db 520 ATGGAGAAGAAATCCAGCATATTTAGCATGGAAGCAACCTCATGACACAGAGCAT 579
Qy 140 PheGlnAsnPheSerIleThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsn 159
Db 580 TTCCAAATTTTCAGATGACACTGATCAAGATTTAATGATATTAATCTCTCGACGTAAGT 639
Qy 160 SerLeuLeuSerSerIleGlnGluHisGluAsnIleGlyAspIleSerLysSerLeu 179
Db 640 ACCTTGTTTCTCAGTCAGGACATGGGAATGCAATKATKATGAAATCTCAAGTCTCTTA 699
Qy 180 ValGlyLeuAsnThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArg 199
Db 700 ATAAAGTTGAATACCATCTTCTGATTTGAGCTCAACATAGAAAATCTGAATGGCAAA 759
Qy 200 ValGlnGluAsnAlaPheLysGlnGlnGluMetArgLysLeuGluGluArgIleTyr 219
Db 760 ATCCAGAGAAATACCTTCAACAAACAGAGAAATCAGTAATTAGAGGCGGTGTAC 819
Qy 220 AsnAlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyrLeuGlnGlnGluIle 239

```

```

Db 820 AATGTATCGCAGAAATATATGGCTATGAAGAGACAGAGTGCATTTTGGAAACAGAAATA 879
Qy 240 LysGlyGluMetLysLeuLeuAsnAsnIleThrAsnAspLeuArgLeuLysAspTyrGlu 259
Db 880 AAGGAGAAGTGAAGTACTGATTAACATCACTAATGATCTCAGACTGAAAGATTGGAA 939
Qy 260 HisSerGlnThrLeuLysAsnIleThrLeuLeuGlnGly 272
Db 940 CATTTCTCAGACCTTGAAGAAATATCAATTTAATCAAGT 978

RESULT 4
US-10-995-561-371
; Sequence 371, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 371
; LENGTH: 3719
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-371

Alignment Scores:
Pred. No.: 3,25e-101 Length: 3719
Score: 1016.50 Matches: 193
Percent Similarity: 85.3% Conservative: 40
Best Local Similarity: 70.7% Mismatches: 39
Query Match: 49.4% Indels: 1
DB: 8 Gaps: 1

US-10-618-570-2 (1-400) x US-10-995-561-371 (1-3719)

Qy 1 MetAlaGlnTTPAspAAspPheProAspGlnGlnAspThrAspSerCysThrGluSer 20
Db 175 ATGGAGCAGTGGGATCACTTTTCACATCAACAGAGGACACTGATAGCTGCGAATCT 234
Qy 21 ValLysPheAspAlaArgSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
Db 235 GTGAAATTTGATGCTCGCTCAATGACAGCTTTGCTTCCTCGAATCTCAAAACAGCCCT 294
Qy 41 ThrLeuGlnGluArgMetLysSerTyrLysThrAlaLeuLeuThrLeuTyrLeuVal 60
Db 295 TMCCTTCAAGAGAARCTCAAGTCTTCAAAAGCTGCACCTGRTTGCCTTTACCTCTCGTG 354
Qy 61 PheValValLeuValProIleGlyLeuValAlaAlaGlnLeuLeuLysThrGluThr 80
Db 355 TTTGAGTTCTCACTCCCTCTCATTTGGAATAGTGGAGCTCAACTCTCGAATGGGAACG 414
Qy 81 LysAsnCysThrValGlySerValAsnAla---AspIleSerProSerProGluGlyLys 99
Db 415 AAGATTGCTCAGTTAGTTCAACTAATGAAATGATATTAATCAAGTCTCAACGGAAA 474
Qy 100 GlyAsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsn 119
Db 475 GGAATGACACGAAAGAGAAATGAGATTTCAAGAAAGCTTTTATGGAACACATGAGCAAC 534
Qy 120 MetGluSerArgIleGlnTyrLeuSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsn 139
Db 535 ATGGAGAAGAAATCCAGCATATTTAGACATGGAAGCAACCTCATGAGCAGAGCAT 594
Qy 140 PheGlnAsnPheSerIleThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsn 159
Db 595 TTCCAAATTTTCAGCATGACAACTGATCAAGATTTAATGACATTTCTTCGACGTAAGT 654

```

```
Qy 160 SerLeuLeuSerSerIleGlnGluHisGluAenIleIleGlyAspIleSerLysSerLeu 179
    ::::::::::::::::::::
Db 655 ACCITGTTTCTCAGTCAGGACATGGGAATGCAATAKATGAAATCTCCAGTCCCTTA 714

Qy 180 ValGlyLeuAenThrValLeuAspLeuGlnPheSerIleGluThrLeuAenGlyArg 199
    ::::::::::::::::::::
Db 715 ATAAGTTTGAATACCACTGCTGATTTGCAGCTCAACATAGAAAATCTGAATGGCAA 774

Qy 200 ValGlnGluAenAlaPheIleGlnGluMetArgLysLeuGluGluArgIleTyr 219
    ::::::::::::::::::::
Db 775 ATCCAAGAGAATACCTTCAACAAACACAGAGAAATCAGTAATATAGAGGCGGTGTAC 834

Qy 220 AsnAlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyrLeuGluGlnGluIle 239
    ::::::::::::::::::::
Db 835 AATGTATCRGAGAAATATGCTATGAAGAAGAACAGTGCATTTGGACAGGAATA 894

Qy 240 LysGlyGluMetLysLeuAenAenIleThrAenAspLeuArgLeuLysAspTrpGlu 259
    ::::::::::::::::::::
Db 895 AAAGAGAAGTCAAGTACTGAATAACATCACTAATGATCTCAGACTGAAGAATGGAA 954

Qy 260 HisSerGlnThrLeuLysAenIleThrLeuLeuGlnGly 272
Db 955 CATCTCAGACCTTGAGAAATATCAATTAATCAAGT 993

RESULT 5
US-10-995-561-368
; Sequence 368, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 368
; LENGTH: 3791
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-368

Alignment Scores:
Pred. No.: 3,33e-101 Length: 3791
Score: 1016.50 Matches: 193
Percent Similarity: 85.3% Conservative: 40
Best Local Similarity: 70.7% Mismatches: 39
Query Match: 49.4% Indels: 1
DB: 8 Gaps: 1

US-10-618-570-2 (1-400) x US-10-995-561-368 (1-3791)

Qy 1 MetAlaGlnTrpAspPheProAspGlnGlnGluAspThrAspSerCysThrGluSer 20
Db 160 ATGGAGCAGTGGATCACTTTCAATCAACAGGAGGACATGATAGCTCTCGAATCT 219

Qy 21 ValLysPheAspAlaArgSerValThrAlaLeuLeuProHisProLysAenGlyPro 40
Db 220 GTGAATTTGATCGCTCAATGACAGCTTTGCTTCCTCGAATSCCTAAACACAGCCT 279

Qy 41 ThrLeuGlnGluArgMetLysSerTyrLysThrAlaLeuIleThrLeuTyrLeuIleVal 60
Db 280 TWCCTTCAAGAGAARCTGAAGTCTTCAAAAGTGCACCTGRTTGGCCCTTACCTCTCGTG 339

Qy 61 PheValValLeuValProIleGlyIleValAlaAlaGlnLeuLeuLysTyrTrpLthr 80
Db 340 TTTGAGTTCATCCCTCATGGAATAGTGGAGCTCACTCCTGAGTGGGAACG 399

Qy 81 LysAenCysThrValGlySerValAenAla---AspIleSerProSerProGluGlyLys 99
Db 400 AAGAATTCCTCAGTTAGTTCAACTAATGCAAAATGATATAACTCAAAAGTCTCACGGGAAA 459
```

```
Qy 100 GlyAenGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAen 119
    ::::::::::::::::::::
Db 460 GGAATATCAGACGAAAGAGAAATGAGATTTCAAGAGYCTTTATGGAAACACATAGGCAAC 519

Qy 120 MetGluSerArgIleGlnTyrLeuSerAspAenGluAlaAenLeuLeuAspAlaLysAen 139
    ::::::::::::::::::::
Db 520 ATGGAGAAGAGATCCAGCATATTTTAGACATGATGAAGCAACCTCATGGACACAGAGCAT 579

Qy 140 PheGlnAenPheSerIleThrThrAspGlnArgPheAenAspValLeuPheGlnLeuAen 159
    ::::::::::::::::::::
Db 580 TTCCAAAATTTACGATGACCAACTGATCAAGATTTAATGACATTTCTTCGACGTAAAT 639

Qy 160 SerLeuLeuSerSerIleGlnGluHisGluAenIleIleGlyAspIleSerLysSerLeu 179
    ::::::::::::::::::::
Db 640 ACCTTGTTTCTCAGTCCAGGACATGGGATGCAATATAGAAATCTCCAGTCCCTTA 699

Qy 180 ValGlyLeuAenThrValLeuAspLeuGlnPheSerIleGluThrLeuAenGlyArg 199
    ::::::::::::::::::::
Db 700 ATAAGTTTGAATACCACTGCTTGTGTTGCTGAGCTCAACATAGAAAATCTGAATGGCAA 759

Qy 200 ValGlnGluAenAlaPheLysGlnGlnGluMetArgLysLeuGluGluArgIleTyr 219
    ::::::::::::::::::::
Db 760 ATCCAAGAGAATACCTTCAAAACACAAAGAGAAATCAGTAATATAGAGGAGCGTGTTC 819

Qy 220 AsnAlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyrLeuGluGlnGluIle 239
    ::::::::::::::::::::
Db 820 AATGTATCRGAGAAATATGCTATGAAGAAGAACAGTGCATTTGGAAACAGGAATA 879

Qy 240 LysGlyGluMetLysLeuAenAenIleThrAenAspLeuArgLeuLysAspTrpGlu 259
    ::::::::::::::::::::
Db 880 AAAGAGAAGTGAAGTACTGAATAACATCACTAATGATCTCAGACTGAAGAATGGAA 939

Qy 260 HisSerGlnThrLeuLysAenIleThrLeuLeuGlnGly 272
Db 940 CATCTCAGACCTTGAGAAATATCAATTAATCAAGT 978

RESULT 6
US-10-995-561-13314
; Sequence 13314, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13314
; LENGTH: 119036
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(119036)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-995-561-13314

Alignment Scores:
Pred. No.: 3.19e-41 Length: 119036
Score: 490.50 Matches: 121
Percent Similarity: 56.7% Conservative: 45
Best Local Similarity: 41.3% Mismatches: 78
Query Match: 23.8% Indels: 49
DB: 8 Gaps: 8

US-10-618-570-2 (1-400) x US-10-995-561-13314 (1-119036)

Qy 37 LysAenGlyProThrLeuGlnGluArgMetLysSerTyrLysThrAlaLeuIleThr--- 55
    ::::::::::::::::::::
```


Db 52113 CGGATGGAATGATGGGATCCTGTATGCTTCAAGGATCAGCCATGACTGGT 52172
Qy 56 LeuTyrLeuIleValPheValValLeuValProIleIleGlyIleValAlaAlaGlnLeu 75
Db 52173 TTAATTTTAACTAAATAT-----ACAGCTCAACTC 52202
Qy 76 LeuLysTrpGluThrLysAsnCythrValGlySerValAsnAla---AspIleSerPro 94
Db 52203 CTGAAGTGGGAACCAAGAAATGCTCAGTTAGTTCACTAATGCAATATGATTAACCTAA 52262
Qy 95 SerProGluGlyLysGlyAsnGlySerGluAspGluMetArgPheArgGluAlaValMet 114
Db 52263 AGTCTCAGCGGAAAGGAATACACAGCRAAGGAAGAAATGAGATTTCAAGAAGCTTTATG 52322
Qy 115 GluArgMetSerAsnMetGluSerArgIleGlnTyrLeuSerAspAsnGluAlaAsnLeu 134
Db 52323 GAACACATGAGCAACATGAGAGAGAAATCCAGCATATTTTAGACATGGAAGCCAACTC 52382
Qy 135 LeuAspAlaLysAsnPheGlnAsnPheSerIleThrThrAspGlnArgPheAsnAspVal 154
Db 52383 ATGGACACAGAGCATTTCCAAATTTTCAGCATGACACTGATCAAGATTTAATGACATT 52442
Qy 155 LeuPheGlnLeuAsnSerLeuSerSerIleGlnGluHisGluAsnIleIleGlyAsp 174
Db 52443 CTCTCGCAGCTAAGTACCTTGTCTCAGTCAGGACATGGGAATGCAATATAGTAA 52502
Qy 175 IleSerLysSerLeuValGlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGlu 194
Db 52503 ATCTCCAAGTCTTAATAGTTGAATACCACTATGCTGTATTTGTCAGCTCAACATAGAA 52562
Qy 195 ThrLeuAsnGlyArgValGluAsnAlaPheLysGlnGluAsnMetArgLysLeu 214
Db 52563 AATCTGAATGCAAAATCCAGAGAAATACCTTCAACAAACAA----- 52604
Qy 215 GluGluArgIleTyrAsnAlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyr 234
Db 52605 GAGGTAGAGTTTGGAGTCAAGTCACTGTGGAAGTCACTCACTCAAGTCTCGT 52664
Qy 235 LeuGluGlnGluIleLysGlyGluMetLysLeuLeuAsnAsnIleThrAsnAsp----- 252
Db 52665 CTCCTTAGCGAGTACAGG---AAAGCAAAAGTTGCAAACTGTTTCATCATCTGTTTA 52721
Qy 253 -----LeuArgLeuLysAspTrpGluHisSer---GlnThrLeuLysAsn 266
Db 52722 CCCTGATTTCTATGCTCACTATTCAGATTTTAAACGAATGTTAAGGATTTAAATGG 52781
Qy 267 IleThrLeuGlnGlyAlaArgLysCySerLeuThrGlyLysTrpThrAsnAspLeu 286
Db 52782 ATTCGCTTCTCTAAACTCTTAGAAA----- 52808
Qy 287 GlySerAsnMetThrIleGlyAlaValAsnSerArgGlyGluPheThrGlyThrTyrIle 306
Db 52809 -----AGGGGAATTTTCCCATGGAATATAG 52835
Qy 307 ThrAlaValThrAlaThrSerAsnGluIleLysGluSer 319
Db 52836 ACTCCCTTACATATTTTATAAATACTGTTAAATATGCT 52874

RESULT 7

US-10-995-561-9974
; Sequence 9974, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; NUMBER OF SEQ ID NOS: 2004-11-24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9974

; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-9974
Alignment Scores:
Pred. No.: 7.5e-22 Length: 201
Score: 282.00 Matches: 52
Percent Similarity: 93.9% Conservativeness: 10
Best Local Similarity: 78.8% Mismatches: 4
Query Match: 13.7% Indels: 0
DB: 8 Gaps: 0
US-10-618-570-2 (1-400) x US-10-995-561-9974 (1-201)
Qy 13 AspThrAspSerCythrThrGluSerValLysPheAspAlaArgSerValThrAlaLeuLeu 32
Db 3 GACACTGATAGTCTCGGAATCTGTGAAATTTGATGCTCGCTCAATGACAGCTTTGCTT 62
Qy 33 ProProHisProLysAsnGlyProThrLeuGlnGluArgMetLysSerTyrLysThrAla 52
Db 63 CCTCCGAATCCTAAAAACAGCCCTTCCCTTCAAGAGAARCTGAAAGTCTTCAAGAGCTGCA 122
Qy 53 LeuIleThrLeuTyrLeuIleValPheValValLeuValProIleIleGlyIleValAla 72
Db 123 CTGATTCCTTTACCTCCTCGTGTGTTGCGAGTTCTCATCCCTCTCATTTGGAAATAGTGCA 182
Qy 73 AlaGlnLeuLeuLysTrp 78
Db 183 GCTCAACTCCTCGAAGTGG 200
RESULT 8
US-10-995-561-10011
; Sequence 10011, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10011
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-10011
Alignment Scores:
Pred. No.: 7.5e-22 Length: 201
Score: 282.00 Matches: 52
Percent Similarity: 93.9% Conservativeness: 10
Best Local Similarity: 78.8% Mismatches: 4
Query Match: 13.7% Indels: 0
DB: 8 Gaps: 0
US-10-618-570-2 (1-400) x US-10-995-561-10011 (1-201)
Qy 13 AspThrAspSerCythrThrGluSerValLysPheAspAlaArgSerValThrAlaLeuLeu 32
Db 3 GACACTGATAGTCTCGGAATCTGTGAAATTTGATGCTCGCTCAATGACAGCTTTGCTT 62
Qy 33 ProProHisProLysAsnGlyProThrLeuGlnGluArgMetLysSerTyrLysThrAla 52
Db 63 CCTCCGAATCCTAAAAACAGCCCTTCCCTTCAAGAGAARCTGAAAGTCTTCAAGAGCTGCA 122
Qy 53 LeuIleThrLeuTyrLeuIleValPheValValLeuValProIleIleGlyIleValAla 72
Db 123 CTGATTCCTTTACCTCCTCGTGTGTTGCGAGTTCTCATCCCTCTCATTTGGAAATAGTGCA 182


```

; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9940
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-9940

Alignment Scores:
Pred. No.: 7,39e-21 Length: 201
Score: 273.00 Matches: 51
Percent Similarity: 92.4% Conservative: 10
Best Local Similarity: 77.3% Mismatches: 5
Query Match: 13.3% Indels: 0
Gaps: 0

US-10-618-570-2 (1-400) x US-10-995-561-9940 (1-201)

Qy 9 AspGlnGlnGluAspThrAspSerCysThrGluSerValPheAspAlaArgSerVal 28
Db 4 AATCAACAGGAGGACACTGATGCTCCGAATCTGTGAATTTGATGCTCGCTCAATG 63
Qy 29 ThrAlaLeuLeuProHisProLysAsnGlyProThrLeuGlnGluArgMetLysSer 48
Db 64 ACAGCTTTGCTCTCCGAATCTTAAACACAGCCCTTMCCTTCAAGAGAACTGAAGTCC 123
Qy 49 TyrlsThrAlaLeuLeuThrLeuTyrlsLeuLeuValPheValLeuValProLeuLe 68
Db 124 TTCAAGCTGCACGTGATGGCCCTTTACCTCTCGTGTTCAGTGTCTCATCCCTCTCATT 183
Qy 69 GlyIleValAlaAlaGln 74
Db 184 GGAATAGTGGCAGCTCAA 201

RESULT 14
US-10-995-561-10014
; Sequence 10014, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10014
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-10014

Alignment Scores:
Pred. No.: 7,39e-21 Length: 201
Score: 273.00 Matches: 51
Percent Similarity: 92.4% Conservative: 10
Best Local Similarity: 77.3% Mismatches: 5
Query Match: 13.3% Indels: 0
Gaps: 0

US-10-618-570-2 (1-400) x US-10-995-561-10014 (1-201)

Qy 9 AspGlnGlnGluAspThrAspSerCysThrGluSerValPheAspAlaArgSerVal 28
Db 4 AATCAACAGGAGGACACTGATGCTCCGAATCTGTGAATTTGATGCTCGCTCAATG 63
Qy 29 ThrAlaLeuLeuProHisProLysAsnGlyProThrLeuGlnGluArgMetLysSer 48
Db 64 ACAGCTTTGCTCTCCGAATCTTAAACACAGCCCTTMCCTTCAAGAGAACTGAAGTCC 123
Qy 49 TyrlsThrAlaLeuLeuThrLeuTyrlsLeuLeuValPheValLeuValProLeuLe 68
Db 124 TTCAAGCTGCACGTGATGGCCCTTTACCTCTCGTGTTCAGTGTCTCATCCCTCTCATT 183
Qy 69 GlyIleValAlaAlaGln 74
Db 184 GGAATAGTGGCAGCTCAA 201

RESULT 15
US-10-995-561-10049
; Sequence 10049, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24

```

```
/ NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10049
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-10049

Alignment Scores:
Pred. No.:      7.38e-21      Length:      201
Score:          273.00      Matches:      51
Percent Similarity: 92.4%      Conservative: 10
Best Local Similarity: 77.3%      Mismatches: 5
Query Match:    13.3%      Indels:      0
DB:             8          Gaps:      0

US-10-618-570-2 (1-400) x US-10-995-561-10049 (1-201)

Qy      9 AspGlnGlnGluAspThrAspSerCysThrGluSerValLysPheAspAlaArgSerVal 28
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      4 AATCAACAGGAGGACACTGATAGCTGCTCCGAATCTGGAATTTGATGCTCGCTCAATG 63

Qy     29 ThrAlaLeuLeuProHisProLysAsnGlyProThrLeuGlnGluArgMetLysSer 48
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db     64 ACAGCTTTGCTTCCTCCGAATCTATAAAACAGCCCTTWCCTTCAAGAGAAACTGAAGTCC 123

Qy     49 TyrLysThrAlaLeuIleThrLeuTyrLeuIleValPheValValLeuValProIleIle 68
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db    124 TTCAAGCTGCACCTGATTTGCCCTTTACCTCTCGTGTTCAGTTTCATCCCTCTCATTT 183

Qy     69 GlyIleValAlaAlaGln 74
      ::::::::::::::::::::::
Db    184 GGAATAGTGGCAGCTCAA 201
```

Search completed: February 27, 2006, 09:49:08
Job time : 545 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 27, 2006, 09:22:16 ; Search time 230 Seconds
(without alignments)
3091.412 Million cell updates/sec

Title: US-10-618-570-2

Perfect score: 2059

Sequence: 1 MAQWDDPPDQEDTDSCTES.....DWKATRVGINFTRLRTQKE 400

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/abss/ABSSWEB pool/US10618570/runat 27022006 063859 13138/app query.fasta_1
-DB=Issued Patents NA -QWTF=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPTCL=0
-LOOPTXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-HOST=abss03p -USER=US10618570 @CGN 1 1 290 @runat 27022006 063859 13138
-NCPU=6 -ICPU=3 -NO WMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAEXT=0.5 -FGAPOP=6
-FGAEXT=7 -YGAPOP=10 -YGAEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:

1: /cgn2_6/ptodata/1/ina/1 COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5 COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq:*
5: /cgn2_6/ptodata/1/ina/H COMB.seq:*
6: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq:*
7: /cgn2_6/ptodata/1/ina/PP COMB.seq:*
8: /cgn2_6/ptodata/1/ina/RE COMB.seq:*
9: /cgn2_6/ptodata/1/ina/backfiles.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|----|--------------------|
| 1 | 1381.5 | 67.1 | 1588 | 9 | 5510466-3 |
| 2 | 1381.5 | 67.1 | 1807 | 9 | 5510466-1 |
| 3 | 1042.5 | 50.6 | 1366 | 3 | US-09-949-016-3387 |
| 4 | 1042.5 | 50.6 | 1367 | 2 | US-08-453-117-3 |
| 5 | 1042.5 | 50.6 | 1367 | 2 | US-08-948-222-3 |
| 6 | 1042.5 | 50.6 | 1367 | 2 | US-08-973-145-3 |
| 7 | 1042.5 | 50.6 | 1367 | 3 | US-09-016-434-1260 |
| 8 | 1042.5 | 50.6 | 1367 | 6 | PCT-US96-08081-3 |
| 9 | 1042.5 | 50.6 | 2028 | 2 | US-08-453-117-1 |

| | | | | | | |
|----|--------|------|-------|---|----------------------|--------------------|
| 10 | 1042.5 | 50.6 | 2028 | 2 | US-08-948-222-1 | Sequence 1, Appli |
| 11 | 1042.5 | 50.6 | 2028 | 2 | US-08-973-145-1 | Sequence 1, Appli |
| 12 | 1042.5 | 50.6 | 2028 | 6 | PCT-US96-08081-1 | Sequence 1, Appli |
| 13 | 1039.5 | 50.5 | 2037 | 2 | US-08-154-365-1 | Sequence 1, Appli |
| 14 | 688 | 33.4 | 484 | 2 | US-08-554-586-1 | Sequence 1, Appli |
| 15 | 678 | 32.9 | 604 | 3 | US-08-831-399-3 | Sequence 3, Appli |
| 16 | 678 | 32.9 | 604 | 3 | US-09-366-862-3 | Sequence 3, Appli |
| 17 | 678 | 32.9 | 604 | 3 | US-09-368-772-3 | Sequence 3, Appli |
| 18 | 495.5 | 24.1 | 55886 | 3 | US-09-949-016-15129 | Sequence 15129, A |
| 19 | 275 | 13.4 | 601 | 3 | US-09-949-016-121418 | Sequence 121418, A |
| 20 | 173 | 8.4 | 245 | 3 | US-09-513-999C-36319 | Sequence 36319, A |
| 21 | 156.5 | 7.6 | 601 | 3 | US-09-949-016-121413 | Sequence 121413, A |
| 22 | 147.5 | 7.2 | 1173 | 3 | US-09-142-974B-4 | Sequence 4, Appli |
| 23 | 145.5 | 7.1 | 1176 | 3 | US-09-142-974B-3 | Sequence 3, Appli |
| 24 | 138.5 | 6.7 | 498 | 3 | US-09-117-447-7 | Sequence 7, Appli |
| 25 | 138.5 | 6.7 | 525 | 2 | US-07-924-028A-2 | Sequence 2, Appli |
| 26 | 138.5 | 6.7 | 552 | 6 | PCT-US93-05240-13 | Sequence 13, Appli |
| 27 | 138.5 | 6.7 | 638 | 3 | US-08-831-399-1 | Sequence 1, Appli |
| 28 | 138.5 | 6.7 | 638 | 3 | US-08-381-430-1 | Sequence 1, Appli |
| 29 | 138.5 | 6.7 | 638 | 3 | US-09-366-862-1 | Sequence 1, Appli |
| 30 | 138.5 | 6.7 | 638 | 3 | US-09-368-772-1 | Sequence 1, Appli |
| 31 | 138.5 | 6.7 | 1131 | 9 | 5168049-1 | Patent No. 5168049 |
| 32 | 138.5 | 6.7 | 1266 | 3 | US-09-938-270B-2 | Sequence 2, Appli |
| 33 | 138.5 | 6.7 | 1296 | 2 | US-08-491-988-6 | Sequence 6, Appli |
| 34 | 138.5 | 6.7 | 1356 | 2 | US-08-491-988-4 | Sequence 4, Appli |
| 35 | 135.5 | 6.6 | 384 | 3 | US-08-831-399-15 | Sequence 15, Appli |
| 36 | 135.5 | 6.6 | 384 | 3 | US-09-366-862-15 | Sequence 15, Appli |
| 37 | 135.5 | 6.6 | 384 | 3 | US-09-368-772-15 | Sequence 15, Appli |
| 38 | 135.5 | 6.6 | 387 | 2 | US-08-211-833-1 | Sequence 1, Appli |
| 39 | 135.5 | 6.6 | 387 | 2 | US-08-434-718-1 | Sequence 1, Appli |
| 40 | 135.5 | 6.6 | 1257 | 2 | US-08-491-988-8 | Sequence 8, Appli |
| 41 | 135.5 | 6.6 | 30549 | 3 | US-09-134-001C-322 | Sequence 322, App |
| 42 | 132.5 | 6.4 | 354 | 3 | US-07-780-717C-6 | Sequence 6, Appli |
| 43 | 129.5 | 6.3 | 4407 | 3 | US-09-949-016-1690 | Sequence 1690, Ap |
| 44 | 128.5 | 6.2 | 7425 | 3 | US-09-453-702B-212 | Sequence 212, App |
| 45 | 128.5 | 6.2 | 7425 | 3 | US-10-114-170-212 | Sequence 212, App |

ALIGNMENTS

RESULT 1
5510466-3
; Patent No. 5510466
; APPLICANT: KREIGER, MONTY; KODAMA, TATSUHIKO
; TITLE OF INVENTION: SCAVENGER RECEPTOR PROTEIN AND ANTIBODY
; THERETO
; NUMBER OF SEQUENCES: 12
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/307,400
; FILING DATE: 16-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 997,113
; FILING DATE: 24-DEC-1992
; APPLICATION NUMBER: 391,486
; FILING DATE: 09-AUG-1989
; APPLICATION NUMBER: 272,002
; FILING DATE: 15-NOV-1988
; SEQ ID NO:3:
; LENGTH: 1588
5510466-3

Alignment Scores:
Pred. No.: 4.63e-154 Length: 1588
Score: 1381.50 Matches: 307
Percent Similarity: 79.4% Conservatives: 24
Best Local Similarity: 73.8% Mismatches: 66
Query Match: 9 Indels: 22
DB: Gaps: 6

US-10-618-570-2 (1-400) x 5510466-3 (1-1588)

Qy 1 MetalaGlnTrpAspAppPheProAspGlnGluAspThrAspSerCysThrGluSer 20
|||||

```
Db 4 ATGGCACAGTGGGATGACTTTCTCTGATCAGCAAGAGGACACTGACAGCTGTACAGAGTCT 63
Qy 21 VallysPheAspAlaAargSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
Db 64 GTGAAGTTCGATGCTCGCTCAGTGACAGCTTTGCTCTCCCATCTCTAAATAATGGCCCA 123
Qy 41 ThrLeuGlnGluAargMetLysSerTyrLysThrAlaLeuLeuThrLeuLysLeuVal 60
Db 124 ACTCTTCAAGAGAGGATGAGTCTTATATAAAGTGCACCTGATCACCCTTTATCTCATTTG 183
Qy 61 PheValValLeuValProIleIleGlyIleValAlaAlaGlnLeuLeuLysTyrPgluThr 80
Db 184 TTTGTAGTTCTCGTCCCATCTTGGCATAGTGGCAGCTCAGCTCTCTGAAATGGGAACG 243
Qy 81 LysAsnCysThrValGlySerValAsnAlaAspIleSerProSerProGluGlyGly 100
Db 244 AAGAATTCCAGCGTGGCTCAGTTAATGCAGATATATCTCAAGTCCGGAAGGCCAAGGA 303
Qy 101 AsnGlySerGluAaspGluMetArgPheArgGluAlaValMetGluAargMetSerAsnMet 120
Db 304 AATGGCAGTGAAGATGAATGAGATTTCCAGAGAGCTGTGATGGAAAGCATGAGCAACATG 363
Qy 121 GluSerArgIleGlnTyrLeuSerAspAsnGluAlaAsnLeuLeuAaspAlaLysAsnPhe 140
Db 364 GAAAGCAATCCAGTATCTTTTCAGATAATGAAGCCAAATCTCTAGATGCTAAGAAATTC 423
Qy 141 GlnAsnPheSerIleThrThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsnSer 160
Db 424 CAAAATTTCCAGCATAACTGATCAAGATTTAATGATGTTCTTTTCCAGCTAAATTC 483
Qy 161 LeuLeuSerSerIleGlnGluHisGluAsnIleIleGlyAspIleSerLysSerLeuVal 180
Db 484 TTACTTTCTCCATCCAGAAACATGAGAAATATCATAGGGATATCTCCAGTCAATTAGTA 543
Qy 181 GlyLeuAsnThrThrValLeuAaspLeuGlnPheSerIleGluThrLeuAsnGlyArgVal 200
Db 544 GGTCTGAACACCAACAGTACTTGTATGTCAGTTCAGTATTGAACACACTGAATGGCAGAGTC 603
Qy 201 GlnGluAsnAlaPheLysGlnGlnGluMetArgLysLeuGluGluArgIleTyrAsn 220
Db 604 CAAGAGATGCAATTTAAACAAACAGAGAGAGATGCGTAAATTAGAGGAGCGTATATACAA 663
Qy 221 AlaSerAlaGluIleLysSerLeuAaspGluLysGlnValTyrLeuGluGlnGluIleLys 240
Db 664 GCATCAGCAGAAATTAAGTCTCTAGATGAAACAAACAGTATATTGGACACAGAAATAAAA 723
Qy 241 GlyGluMetLysLeuLeuAsnAsnIleThrAsnAspLeuAargLeuLysAspTyrPgluHis 260
Db 724 GGGGAAATGAAACTGTTCAATAATATCACTAATGATCTGAGGCTGAAGGATTGGGAACAT 783
Qy 261 SerGlnThrLeuLysAsnIleThrLeuLeuGlnGlyAla----- 273
Db 784 TCTCAGACATTTGAAATAATATCACTTTACTTCAAGGTCCTCTCGGACCTCCAGGTGAAAAA 843
Qy 274 -----ArgLysCysSerLeuThrGlyLysIleThrAsnAspLeuGly---SerAsnMet 290
Db 844 GGAGATAGAGGCCCTCTCTGGACA---AATGGTATACAGGCTTTCCAGCTCTAATAGG 899
Qy 291 ThrIleGlyAlaValAsnSerArgGlyGluPheThrGlyThrTyrIleThrAlaValThr 310
Db 900 TACTCCAGTCTTAAAGGTGATCGGGGATCTCTGTTTACTCTGGAGTTCGAGGATCC 959
Qy 311 AlaThrSerAsnGluIleLysGluSer-ProLeuHisGlyThrGlnAsnThrIleAsnLys 330
Db 960 AGGACCAATGGGAAGACCGGAGCCAGGACTTAATGGACAAAAAAGCCAGAGGGGAGA 1019
Qy 330 eArgThrGlnProThrPhe-----GlyPheThrVal-AsnTyrLysPheSerGlu 347
Db 1020 AAAACCCACTCAACCATCCAAACCAATCTTAATACATCCAGTCTGGTGGGTGGCAGCGG 1079
Qy 347 erThrThrValPheThrGlyGlnCysPheIleAspArgAsnGlyLysGluValLeuLysT 367
Db 1080 CCCTCACGAAGGCAGAGTGGAGATTTTTCACGAAGCCAGTGGGTGCGGTGTGTGACGA 1139
```

```
Qy 367 hrMetTyrLeuLeuAargSerVal-AsnAspIleGlyAasp---AspTyrLysAlaThr 385
Db 1140 CCGC-TGGAACTGCGTGGAGGACTGCTGCTCGCAGGAGCTTGGATACAAAGTGTTC 1198
Qy 386 ArgValGlyIleAsn-----IlePheThrArgLeuLeuArgThrGln 398
Db 1199 AAGGTGTGATAGCGAGCTTATTTTGGAAAAAGGTACGGGTCCAA 1243

RESULT 2
5510466-1
; Patent No. 5510466
; APPLICANT: KREIGER, MONTY;KODAMA, TATSUHIKO
; TITLE OF INVENTION: SCAVENGER RECEPTOR PROTEIN AND ANTIBODY
; THEREETO
; NUMBER OF SEQUENCES: 12
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/307,400
; FILING DATE: 16-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 997,113
; FILING DATE: 24-DEC-1992
; APPLICATION NUMBER: 391,486
; FILING DATE: 09-AUG-1989
; APPLICATION NUMBER: 272,002
; FILING DATE: 15-NOV-1988
; SEQ ID NO:1:
; LENGTH: 1807
5510466-1

Alignment Scores:
Pred. No.: 5,76-154 Length: 1807
Score: 1381.50 Matches: 288
Percent Similarity: 85.2% Conservative: 16
Best Local Similarity: 80.7% Mismatches: 34
Query Match: 67.1% Indels: 20
DB: 9 Gaps: 4

US-10-618-570-2 (1-400) x 5510466-1 (1-1807)
Qy 1 MetAlaGlnTyrAspAspPheProApgGlnGlnAspThrAspSerCysThrGluSer 20
Db 36 ATGGCACAGTGGGATGACTTTCTCTGATCAGCAAGAGGACACTGACAGCTGTACAGAGTCT 95
Qy 21 VallysPheAspAlaAargSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
Db 96 GTGAAGTTCGATGCTCGCTCAGTGACAGCTTTGCTCTCCCATCTCTAAATAATGGCCCA 155
Qy 41 ThrLeuGlnGluAargMetLysSerTyrLysThrAlaLeuLeuThrLeuLysLeuVal 60
Db 156 ACTCTTCAAGAGAGGATGAGTCTTATATAAAGTGCACCTGATCACCCTTTATCTCATTTG 215
Qy 61 PheValValLeuValProIleIleGlyIleValAlaAlaGlnLeuLeuLysTyrPgluThr 80
Db 216 TTTGTAGTTCTCGTGGCCCATCTTGGCATAGTGGCAGCTCAGCTCTCTGAAATGGGAACG 275
Qy 81 LysAsnCysThrValGlySerValAsnAlaAspIleSerProSerProGluGlyGly 100
Db 276 AAGAATTTCAGCGTGGCTCAGTTAATGCAGATATATCTCCAGTCCGGAAGGCCAAGGA 335
Qy 101 AsnGlySerGluAaspGluMetArgPheArgGluAlaValMetGluAargMetSerAsnMet 120
Db 336 AATGGCAGTGAAGATGAATGAGATTTCCAGAGAGCTGTGATGGAAAGCATGAGCAACATG 395
Qy 121 GluSerArgIleGlnTyrLeuSerAspAsnGluAlaAsnLeuLeuAaspAlaLysAsnPhe 140
Db 396 GAAAGCAATCCAGTATCTTTTCAGATAATGAAGCCAAATCTCTAGATGCTAAGAAATTC 455
Qy 141 GlnAsnPheSerIleThrThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsnSer 160
Db 456 CAAAATTTCCAGCATAACTGATCAAGATTTAATGATGTTCTTTTCCAGCTAAATTC 515
Qy 161 LeuLeuSerSerIleGlnGluHisGluAsnIleIleGlyAspIleSerLysSerLeuVal 180
```


Db 516 TTACTTCTCCATCCAGAACATGAGATATCATAGGGATATCTCCAGTCAATAGTA 575
Qy GlyLeuAenThrThrValLeuAenGlnPheSerIleGluThrLeuAenGlyArgVal 200
Db 576 GGTCTGAACACACACAGTACTTGAATTCAGTTCAGTATTTGAACACTGAATGGCAGATC 635
Qy GlnGluAenAlaPheLeuGlnGluGluMetArgLysLeuGluGluArgIleThrAen 220
Db 636 CAAGAGATGCATTTAAACAAACAGAGAGATGGTAAATAGAGGACGCTATATACAT 695
Qy AlaSerAlaGluIleLysSerLeuAenGlnValThrLeuGluGlnGluIleLys 240
Db 696 GCATCAGCAGAAATTAAGTCTCTAGATGAAACAAAGTATATTTGGAACAGGAATATAA 755
Qy GlyGluMetLysLeuLeuAenAenIleThrAenAenLeuAenGluAenThrTrpGluHis 260
Db 756 GGGGAAATGAACATGTTGAATTAATATCACTAATGATCTGAGCTGAAGGATGGGAACAT 815
Qy 261 SerGlnThrLeuLysAenIleThrLeuGlnGlnGlyAla----- 273
Db 816 TCTCAGACATTTGAAATAATATCACTTTACTTCCAAAGGTCTCTCGGACCTCCAGGTCAAAAA 875
Qy 274 -----ArgLysCysSerLeuThrGlyLysTrpThrAenAenLeuGly---SerAenMet 290
Db 876 GGAGATGAGGCC-TCTCGACA---AAATGGTATATACAGGCTTCCAGGTCTAATAGG 931
Qy 291 ThrIleGlyAlaValAenSerArgGlyGluPheThrGlyThrThrIleThrAlaValThr 310
Db 932 TACTCCAGGTCTTAAAGGTGATCGGGGATCTCTGTTTACTCGGAGTTCGAGGATCCCC 991
Qy 311 AlaThrSerAenGluIleLysGluSerProLeuHisGlyThrGlnAenThrIleAenLys 330
Db 992 AGGACCAATGGGAAGACCGGAGCCAGGACTTATTGGACAAAAAGCCAGAGGAGAGA 1051
Qy 331 ArgThrGlnProThrPheGlyPheThrValAenTrpLysPheSerGluSer 347
Db 1052 AAAAGGGAG-----TGAAGCATGCAAGAGACC 1078

RESULT 3

US-09-949-016-3387
; Sequence 3387, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED.
; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3387
; LENGTH: 1366
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-3387

Alignment Scores:
Pred. No.: 7,95e-114 Length: 1366
Score: 1042.50 Matches: 197
Percent Similarity: 87.2% Conservative: 41
Best Local Similarity: 72.2% Mismatches: 34
Query Match: 50.6% Indels: 1
DB: 3 Gaps: 1

US-10-618-570-2 (1-400) x US-09-949-016-3387 (1-1366)

Qy 1 MetAlaGlnTrpAspPheProaspGlnGlnGluaspThrAspSerCysThrGluSer 20
Db 67 ATGAGCAGTGGGATCACTTTTCAATCAACAGGAGGACACTGATAGTGTCTCCGAATCT 126
Qy 21 ValLysPheAenAlaAArgSerValThrAlaLeuLeuProProHisProLysAenGlyPro 40
Db 127 GTGAATTTGATGCTCGCTCAATGACACTTTGCTCTCCGAATCTCTCCGAATCTCTCCG 186
Qy 41 ThrLeuGlnGluArgMetLysSerTrpLysThrAlaLeuIleThrLeuTrpLeuIleVal 60
Db 187 TCCCTTCAAGAGAAACGAGTCTTCAAGCTGCACTGATTCCTCTTACCTCTCTCGTG 246
Qy 61 PheValValLeuValProIleGlyLeValAlaGlnLeuLeuLysTrpGluThr 80
Db 247 TTTGAGTTCTCATCCCTCTCATTTGGAATAGTGGCAGCTCAACTCTCTGAAGTGGGAACG 306
Qy 81 LysAenCysThrValGlySerValAenAla---AspIleSerProSerProGluGlyLys 99
Db 307 AAGAAATGCTCAGTTAGTTCAACTTAATGCAATGATATACTCAAGTCTTCACGGGAAAA 366
Qy 100 GlyAenGlySerGluAenGluMetArgPheArgGluAlaValMetGluArgMetSerAen 119
Db 367 GGAATGACAGGAGGAGGAATGAGATTTCAAGAGCTTTTATGGAACACATGAGCAAC 426
Qy 120 MetGluSerArgIleGlnThrLysSerAenGluAlaAenLeuLeuAenAlaLysAen 139
Db 427 ATGGAAGAAGAGATCCAGCATATTTAGACATGGAAGCAACCTCATGACACACAGCAT 486
Qy 140 PheGlnAenPheSerIleThrAspGlnArgPheAenAenValLeuPheGlnLeuAen 159
Db 487 TTCCAAATTTTCAAGATGACAACTGATCAAAAGATTTAATGACATTTCTTCTCGAGCTAAGT 546
Qy 160 SerLeuLeuSerSerIleGlnGluHisGluAenIleLeuGlyAenPheLysSerLysLeu 179
Db 547 ACCTTGTCTTCTCAGTCCAGGACATGGGATGCAATAGATGAATCTCCAGTCTCTTA 606
Qy 180 ValGlyLeuAenThrThrValLeuAenGlnPheSerIleGluThrLeuAenGlyArg 199
Db 607 ATAAGTTTGAATACCATGCTTGTGTTGAGTCAACATAGATAAATCTGAATGGCAAA 666
Qy 200 ValGlnGluAenAlaPheLysGlnGluGluMetArgLysLeuGluGluArgIleThr 219
Db 667 ATCCAAAGAGATACCTTCAAAACAAAGAGAAATCAGTAAATTAGAGGCGGTGTTTAC 726
Qy 220 AsnAlaSerAlaGluIleLysSerLeuAenGlnLysGlnValThrLeuGluGlnGluIle 239
Db 727 AATGATCAGCAGAAATTTATGCTATGAAAGAGAAACAAAGTGCATTTTGGACAGGAATA 786
Qy 240 LysGlyGluMetLysLeuAenAenIleThrAenAenLeuAenGluAenGluAenTrpGlu 259
Db 787 AAGGAGAGAGTGAAGTACTGAATAACATCACTAATGATCTCAGACTCAAGAGATTGGCA 846
Qy 260 HisSerGlnThrLeuLysAenIleThrLeuGlnGly 272
Db 847 CATTCAGACCTTGAAGAAATATCACTTTAATCAAGGT 885

RESULT 4

US-08-453-117-3
; Sequence 3, Application US/08453117
; Patent No. 5683903
; GENERAL INFORMATION:
; APPLICANT: Lysko, Paul G.
; APPLICANT: Elshourbagy, Nabil A.
; APPLICANT: Branner, Mary E.
; TITLE OF INVENTION: Attachment Enhanced 293 Cells
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham - Corporate Patents
; ADDRESSEE: U.S.
; STREET: Mailcode - UW2220, 709 Swedeland Road
; CITY: King of Prussia
; STATE: Pennsylvania

Pred. No.: 7,96e-114 Length: 1367
 Score: 1042.50 Matches: 197
 Percent Similarity: 87.2% Conservative: 41
 Best Local Similarity: 72.2% Mismatches: 34
 Query Match: 50.6% Indels: 1
 DB: 2 Gaps: 1

US-10-618-570-2 (1-400) x US-08-948-222-3 (1-1367)

Qy 1 MetAlaGlnTrpAspPheProAspGlnGlnAspThrAspSerCysThrGluSer 20
 Db 67 ATGGAGCAGTGGGATCACTTTTCAATCAACAGAGGACACTGATAGCTGCTCCGAATCT 126
 Qy 21 VallysPheAspAlaAArgSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
 Db 127 GTGAATTTGATGCTCGCTCAATGACAGCTTGTCTCCGAATCTCCAAAGAGCCCT 186
 Qy 41 ThrLeuGlnGluArgMetLysSerTyrLysThrAlaLeuLeuThrLeuLeuVal 60
 Db 187 TCCCTTCAGAGAACTGAGTCTTCAGAGTCACTGATGCTGCTTACCTCTCGTG 246
 Qy 61 PheValValLeuValProIleGlyLeValAlaAlaGlnLeuLeuLysTrpGluThr 80
 Db 247 TTTGAGTTCTCATCTCTCTCAATGAGTCTCACTCACTCACTCACTCACTCACT 306
 Qy 81 LysAsnCysThrValGlySerValAsnAla---AspLysSerProSerProGluGlyLys 99
 Db 307 AAGAATTCCTCAGTTAGTTCAGTCACTCACTCACTCACTCACTCACTCACT 366
 Qy 100 GlyAsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsn 119
 Db 367 GGAATTCACAGCGAAGGAATGAGATTTCAAGAGTCTTTATGGACACATGAGCAAC 426
 Qy 120 MetGluSerArgIleGlnTyrLeuSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsn 139
 Db 427 ATGGAGAGAAATCCAGCATATTTAGACATGAGGCACTCATGACACAGAGCAT 486
 Qy 140 PheGlnAsnPheSerIleThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsn 159
 Db 487 TTCCAAATTTTCAAGTCACTGATCAAGATTTAATGACATTTCTTCTGACGTAAGT 546
 Qy 160 SerLeuLeuSerSerIleGlnHisGluAsnIleIleGlyAspIleSerLysSerLeu 179
 Db 547 ACCTTGTCTTCTCAGTCCAGGACATGGGATTCATAGATGAATCTCCAACTCTCTTA 606
 Qy 180 ValGlyLeuAsnThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArg 199
 Db 607 ATAAGTTTGAATACCATCTCTGATTTGAGTCACTCACTCACTCACTCACTCACT 666
 Qy 200 ValGlnGluAsnAlaPheLysGlnGlnGluMetArgLysLeuGluGluArgIleTyr 219
 Db 667 ATCCAAGAGAAATACCTTCAACACACAGAGAAATCAGTAAATTTAGAGGAGCGTGTAC 726
 Qy 220 AsnAlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyrLeuGluGlnGluIle 239
 Db 727 AATGTATCAGCAGAAATTTATGCTTCAAGAGAAACAGAGTGCATTTTGGAAAGAAATA 786
 Qy 240 LysGlyGluMetLysLeuLeuAsnIleThrAsnAspLeuArgLeuLysAspTrpGlu 259
 Db 787 AAGGAGAGTGAAGTACTGTAATACATCACTAATGATCTCAGACTGAAGATTGGAA 846
 Qy 260 HisSerGlnThrLeuLysAsnIleThrLeuLeuGlnGly 272
 Db 847 CATCTCAGACCTTTGAGAAATATCACTTTAATTCAGGT 885

RESULT 6

US-08-973-145-3
 ; Sequence 3, Application US/08973145
 ; Patent No. 5919636
 ; GENERAL INFORMATION:
 ; APPLICANT: Lysko, Paul G.
 ; APPLICANT: Eishourbagy, Nabil A.
 ; APPLICANT: Brawner, Mary E.

TITLE OF INVENTION: Attachment Enhanced 293 Cells
 NUMBER OF SEQUENCES: 4
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: SmithKline Beecham - Corporate Patents U.S.
 STREET: Mailcode - UW2220, 709 Swedeland Road
 CITY: King of Prussia
 STATE: Pennsylvania
 COUNTRY: U.S.A.
 ZIP: 19406-5090
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/973,145
 FILING DATE: 26-NOV-1997
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Elizabeth J. Hecht
 REGISTRATION NUMBER: P-41, 824
 REFERENCE/DOCKET NUMBER: P50338
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (610) 270-5009
 TELEFAX: (610) 270-5090
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1367 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: not relevant
 MOLECULE TYPE: cDNA to mRNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 67...1143
 US-08-973-145-3

Alignment Scores:
 Pred. No.: 7,96e-114 Length: 1367
 Score: 1042.50 Matches: 197
 Percent Similarity: 87.2% Conservative: 41
 Best Local Similarity: 72.2% Mismatches: 34
 Query Match: 50.6% Indels: 1
 DB: 2 Gaps: 1

US-10-618-570-2 (1-400) x US-08-973-145-3 (1-1367)

Qy 1 MetAlaGlnTrpAspPheProAspGlnGlnAspThrAspSerCysThrGluSer 20
 Db 67 ATGGAGCAGTGGGATCACTTTTCAATCAACAGAGGACACTGATAGCTGCTCCGAATCT 126
 Qy 21 VallysPheAspAlaAArgSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
 Db 127 GTGAATTTGATGCTCGCTCAATGACAGCTTGTCTCCGAATCTCCAAAGAGCCCT 186
 Qy 41 ThrLeuGlnGluArgMetLysSerTyrLysThrAlaLeuLeuThrLeuLeuVal 60
 Db 187 TCCCTTCAGAGAACTGAGTCTTCAGAGTCACTGATGCTGCTTACCTCTCGTG 246
 Qy 61 PheValValLeuValProIleGlyLeValAlaAlaGlnLeuLeuLysTrpGluThr 80
 Db 247 TTTGAGTTCTCATCTCTCTCAATGAGTCTCACTCACTCACTCACTCACTCACT 306
 Qy 81 LysAsnCysThrValGlySerValAsnAla---AspLysSerProSerProGluGlyLys 99
 Db 307 AAGAATTCCTCAGTTAGTTCAGTCACTCACTCACTCACTCACTCACTCACT 366
 Qy 100 GlyAsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsn 119
 Db 367 GGAATTCACAGCGAAGGAATGAGATTTCAAGAGTCTTTATGGACACATGAGCAAC 426
 Qy 120 MetGluSerArgIleGlnTyrLeuSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsn 139

Db 427 ATGGAGAGAGAAATCCAGCATATTTAGACATGGAAGCCAACTCATGGACACAGAGCAT 486
 Qy 140 PheGlnAenPheSerIleThrThrAspGlnArgPheAenAspValLeuPheGlnLeuAen 159
 Db 487 TTCCAAAATTTCCAGCATGACAACTGATCAAGATTTAATGACATTTCTTCGAGCTAAGT 546
 Qy 160 SerLeuLeuSerSerIleGlnGlnHisGlnAenIleGlyAspIleSerLysSerLeu 179
 Db 547 ACCTTGTTTCTCAGTCAGGACATGGGAATGCATAGATGAATTCCTCAAGTCCTTA 606
 Qy 180 ValGlyLeuAenThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAenGlyArg 199
 Db 607 ATAAGTTTGAATACCACTGCTTGAATTTGAGCTCAACATAGAAAATCTGAATGGCAAA 666
 Qy 200 ValGlnGlnAenAlaPheGlyGlnGlnGlnMetArgLysLeuGlnGluArgIleTyr 219
 Db 667 ATCCAAAGAGAAATACCTTCAACAAACAGAGAAATACAGTNAATTAGAGGCGTGTTCAC 726
 Qy 220 AsnAlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyrLeuGlnGlnGluIle 239
 Db 727 AATGTATCAGCAGAAATATGCTATGAAAGAGAACAGTGCATTTGGACAGGAATA 786
 Qy 240 LysGlyGluMetLysLeuLeuAenAenIleThrAenAspLeuArgLeuAenAspTrpGlu 259
 Db 787 AAAGGAGAAAGTGAAGTACTGAATAACATCACTAATGATCTCAGACTGAAAGATTGGAA 846
 Qy 260 HisSerGlnThrLeuLysAenIleThrLeuLeuGlnGly 272
 Db 847 CATTTCTCAGACCTTGAGAAATATCATCTTAATTCAGGT 885

RESULT 7

US-09-016-434-1260
 ; Sequence 1260, Application US/09016434
 ; Patent No. 6500938
 ; GENERAL INFORMATION:
 ; APPLICANT: Janice Au-Young
 ; APPLICANT: Jeffrey J. Seilhamer
 ; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
 ; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
 ; NUMBER OF SEQUENCES: 1490
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
 ; STREET: 3174 PORTER DRIVE
 ; CITY: PALO ALTO
 ; STATE: CALIFORNIA
 ; COUNTRY: USA
 ; ZIP: 94304
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/016.434
 ; FILING DATE: HEREWITH
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Zeller, Karen J.
 ; REGISTRATION NUMBER: 37,071
 ; REFERENCE/DOCKET NUMBER: PA-0002 US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (650) 855-0555
 ; TELEFAX: (650) 845-4166
 ; INFORMATION FOR SEQ ID NO: 1260:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1367 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear

IMMEDIATE SOURCE:
 ; LIBRARY: GENBANK
 ; CLONE: G219991
 ; US-09-016-434-1260
 Alignment Scores:
 Pred. No.: 7.96e-114 Length: 1367
 Score: 1042.50 Matches: 197
 Percent Similarity: 87.2% Conservative: 41
 Best Local Similarity: 72.2% Mismatches: 34
 Query Match: 50.6% Indels: 1
 DB: 3 Gaps: 1

US-10-618-570-2 (1-400) x US-09-016-434-1260 (1-1367)

Qy 1 MetAlaGlnTrpAspAspPheProAspGlnGlnAspThrAspSerCysThrGluSer 20
 Db 67 ATGGAGCAGTGGGATCACTTTTCACAATCAACAGAGGACACTGTAGTAGTCTCCGAATCT 126
 Qy 21 ValIysPheAspAlaArgSerValThrAlaLeuLeuProHisProLysAenGlyPro 40
 Db 127 GTGAATTTGATGCTCGCTCAATGACAGCTTTGCTTCTCGAATCTCTAAACACAGCCCT 186
 Qy 41 ThrLeuGlnGluArgMetLysSerTyrLysThrAlaLeuIleThrLeuTyrLeuIleVal 60
 Db 187 TCCCTTCAAGAGAACTGGAAGTCTTCAAGAGTGCACCTGATTCCTTACCTCTCGTG 246
 Qy 61 PheValValLeuValProIleGlyIleValAlaAlaGlnLeuLeuLysTyrGluThr 80
 Db 247 TTTGCAGTTCTCATCCCTCTCATTGGAAATAGTGCAGCTCAACTCTCGAAATGGGAAACG 306
 Qy 81 LysAenCysThrValGlySerValAenAla---AspIleSerProSerProGluGlyLys 99
 Db 307 AAGAATTTGCTCAGTTAGTTCAACTTAATGCAATATATACTCAAGTCTCCACGGGAAA 366
 Qy 100 GlyAenGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAen 119
 Db 367 GGAATACACAGCAGAGAGAAATGAGATTTCAAGAAGCTTTATGGAACACATCAGCAAC 426
 Qy 120 MetGluSerArgIleGlnTyrLeuSerAspAenGluAlaAenLeuLeuAspAlaLysAen 139
 Db 427 ATGGAGAGAGAAATCCAGCATATTTTAGACATGGAAGCCAACTCATGGACACAGAGCAT 486
 Qy 140 PheGlnAenPheSerIleThrThrAspGlnArgPheAenAspValLeuPheGlnLeuAen 159
 Db 487 TTCCAAAATTTCAGCATGACAACTGATCAAGATTTAATGACATTTCTTCGAGCTAAGT 546
 Qy 160 SerLeuLeuSerSerIleGlnGlnHisGlnAenIleIleGlyAspIleSerLysSerLeu 179
 Db 547 ACCTTGTCTTCTCAGTCCAGGACATGGGAATGCATAGATGAATCTCCAACTCCTTA 606
 Qy 180 ValGlyLeuAenThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAenGlyArg 199
 Db 607 ATAAGTTTGAATACCACTGCTTGAATTTGAGCTCAACATAGAAAATCTGAATGGCAAA 666
 Qy 200 ValGlnGlnAenAlaPheGlyGlnGlnGlnMetArgLysLeuGlnGluArgIleTyr 219
 Db 667 ATCCAAAGAGAAATACCTTCAACAAACAGAGAAATACAGTAAATTAGAGGCGTGTTCAC 726
 Qy 220 AsnAlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyrLeuGlnGlnGluIle 239
 Db 727 AATGTATCAGCAGAAATATGCTATGAAAGAGAACAGTGCATTTGGACAGGAATA 786
 Qy 240 LysGlyGluMetLysLeuLeuAenAenIleThrAenAspLeuArgLeuLysAspTrpGlu 259
 Db 787 AAAGGAGAAAGTGAAGTACTGAATAACATCACTAATGATCTCAGACTGAAAGATTGGAA 846
 Qy 260 HisSerGlnThrLeuLysAenIleThrLeuLeuGlnGly 272
 Db 847 CATTTCTCAGACCTTGAGAAATATCATCTTAATTCAGGT 885

RESULT 8

PCT-US96-08081-3

```

; Sequence 3, Application PC/TUS9608081
; GENERAL INFORMATION:
; APPLICANT: SmithKline Beecham Corporation
; TITLE OF INVENTION: Attachment Enhanced 293 Cells
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham - Corporate Patents U.S.
; STREET: Mailcode - UW2220, 709 Swedeland Road
; CITY: King of Prussia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19406-5090
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/08081
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Jervis, Herbert H.
; REGISTRATION NUMBER: 31,171
; REFERENCE/DOCKET NUMBER: P50338
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 270-5019
; TELEFAX: (610) 270-5090
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1367 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: not relevant
; MOLECULE TYPE: cDNA to mRNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 67..1143
; PCT-US96-08081-3

Alignment Scores:
Pred. No.: 7.96e-114 Length: 1367
Score: 1042.50 Matches: 197
Percent Similarity: 87.2% Conservative: 41
Best Local Similarity: 72.2% Mismatches: 34
Query Match: 50.6% Indels: 1
DB: 6 Gaps: 1

US-10-618-570-2 (1-400) x PCT-US96-08081-3 (1-1367)
Qy 1 MetAlaGlnTrpAspPheProAspGlnGlnGluAspThrAspSerCysThrGluSer 20
Db 67 ATGGAGCAGTGGGATCACTTTCACAAATCAACAGAGGACACTGATAGTCTCGCATCT 126
Qy 21 VallysPheAspAlaArgSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
Db 127 GTGAAATTTGATGCTCGCTCAATGACAGCTTGTCTTCGGAATCTTAAACAGCCCT 186
Qy 41 ThrLeuGlnGluArgMetLysSerTyrlsThrAlaLeuLeuThrLeuTyrlleVal 60
Db 187 TCCCTTCAAGAGAAAGTGAAGTCTTCAAGCTGCACCTGATGTCCTTACCTCTCGTG 246
Qy 61 PheValValLeuValProIleGlyLeValAlaAlaGlnLeuLeuLysTrpGluThr 80
Db 247 TTTGCGATTTCTCATCCCTCTCATTTGGATAGTGGCAGCTCAACTCTTGAAGTGGGAACG 306
Qy 81 LysAsnCysThrValGlySerValAsnAla---AspLysSerProSerProGluGlyLys 99
Db 307 AAGAATTGCTCAGTTAGTTCAACTAATGCAATGATATCAAGTCTCAACGGGAAA 366
Qy 100 GlyAsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsn 119
Db 367 GGAATGACAGCGAAGAGGAATGAGATTTCAAGAGCTCTTTATGGAACACATGAGCAAC 426

```

```

Qy 120 MetGluSerArgIleGlnTrpLysSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsn 139
Db 427 ATGGAGAAGAGAAATCCAGCATATTTTAGACATGTAAGCCAACTCATGGAACAGAGCAT 486
Qy 140 PheGlnAsnPheSerIleThrThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsn 159
Db 487 TTCCAAATTTTCAGCATGACAACTGATCAAGATTTTAATGACATCTTCTCGCAGTAAGT 546
Qy 160 SerLeuLeuSerSerIleGlnGluHisGluAsnIleIleGlyAspIleSerLysSerLeu 179
Db 547 ACCTTGTTTTCCTCAGTCCAGGAGACATGGGAATGCAATAGATGAATCTCCAAGTCTTA 606
Qy 180 ValGlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArg 199
Db 607 ATAAATTTGAATACCACTTGTGTTGATTTGAGCTCAACATAGAAAATCTGAATGGCAAA 666
Qy 200 ValGlnGluAsnAlaPheLysGlnGlnGluMetArgLysLeuGluArgIleTyrl 219
Db 667 ATCCAGAGAAATACCTTCAACAAACAAAGAGAAATCAGTAAATTAGAGGCGGTGTTTAC 726
Qy 220 AsnAlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyrLeuGluGlnGluIle 239
Db 727 AATGATCAGCAGAAATATGCTGCTATCAAGAGAACACAGTGCATTTGGACACAGAAATA 786
Qy 240 LysGlyGluMetLysLeuLeuAsnAsnIleThrAsnAspLeuArgLeuLysAspTrpGlu 259
Db 787 AAAGGAGAACTGAAAGTACTGAAATAACATCACTAATGATCTCAGACTGAAAGATTGGAA 846
Qy 260 HisSerGlnThrLeuLysAsnIleThrLeuLeuGlnGly 272
Db 847 CATTCAGACCTTGAGAAATATCACTTTAATTCAGGT 885

```

RESULT 9

```

US-08-453-117-1
; Sequence 1, Application US/08453117
; Patent No. 5683903
; GENERAL INFORMATION:
; APPLICANT: Lysko, Paul G.
; APPLICANT: Elehourbagy, Nabil A.
; APPLICANT: Brawner, Mary E.
; TITLE OF INVENTION: Attachment Enhanced 293 Cells
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham - Corporate Patents
; ADDRESSEE: U.S.
; STREET: Mailcode - UW2220, 709 Swedeland Road
; CITY: King of Prussia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19406-5090
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/453,117
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jervis, Herbert H.
; REGISTRATION NUMBER: 31,171
; REFERENCE/DOCKET NUMBER: SBC-P50338
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 270-5019
; TELEFAX: (610) 270-5090
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2028 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: not relevant

```

```

; MOLECULE TYPE: cDNA to mRNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 47..1402
US-08-453-117-1
Alignment Scores:
Pred. No.: 1,5e-113 Length: 2028
Score: 1042.50 Matches: 197
Percent Similarity: 87.2% Conservativeness: 41
Best Local Similarity: 72.2% Mismatches: 34
Query Match: 50.6% Indels: 1
DB: 2 Gaps: 1

US-10-618-570-2 (1-400) x US-08-453-117-1 (1-2028)
Qy 1 MetAlaGlnTrpAspPheProAspGlnGlnGluAspThrAspSerCysThrGluSer 20
Db 47 ATGGAGCAGTGGGATCACTTTCAACAATCAACAGGAGGACACTGATAGCTGCTCCGAATCT 106
Qy 21 ValLysPheAspAlaAatGSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
Db 107 GTGAAATTTGATGCTCGCTCAATGACAGCTTTGCTTCCTCGAATCTTAAACACAGCCCT 166
Qy 41 ThrLeuGlnGluArgMetLysSerTyxLysThrAlaLeuLeuThrLeuTyxLeuIleVal 60
Db 167 TCCCTTCAAGAGAACTGAGTCTTCAAAAGCTGCACCTGATTTGCCCTTTACTCTCTCGTG 226
Qy 61 PheValValLeuValProIleGlyIleValAlaAlaGlnLeuLeuLeuTyxTrpGluThr 80
Db 227 TTTCAGTCTTCATCCCTCTCATTTGGAATAGTGGCAGCTCAACTCTGAAAGTGGAAACG 286
Qy 81 LysAsnCysThrValGlySerValAsnAla---AspLysSerProSerProGluGlyLys 99
Db 287 AAGAAATTCCTCAGTATGTTCAACTAATGCAATATGATATACTCAAAAGTCTCACGGGAAA 346
Qy 100 GlyAsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsn 119
Db 347 GGAATGACAGGAGAGAGAAATGAGATTTCAAGNAGCTTTTATGGAACACATGAGCAAC 406
Qy 120 MetGluSerArgLleGlnTyxLeuSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsn 139
Db 407 ATGGAGAGAGAAATCCAGCATATTTAGACATGGAAGCAACCTCATGGACACAGAGCAT 466
Qy 140 PheGlnAsnPheSerIleThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsn 159
Db 467 TTCCAAAATTTCAAGCATGACAACTGATCAAAAGATTTAATGACATTTCTTCGACGTAAGT 526
Qy 160 SerLeuLeuSerSerIleGlnGluHisGluAsnIleLeGlyAspIleSerLysSerLeu 179
Db 527 ACCTTGTTTCTCAGTCCAGGACATGGGAATGCAATAGATGAATCTCCAAATCTCCCTTA 586
Qy 180 ValGlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArg 199
Db 587 ATAAGTTTGAATACCAATGCTTGATTTGAGCTCAACATAGAAAATCTTGAATGGCAAA 646
Qy 200 ValGlnGluAsnAlaPheLysGlnGlnGlnGluMetArgLysLeuGluGluArgIleTyx 219
Db 647 ATCCAAGAGAAATACCTTCAAAACACAGAGAAATACAGTAAATTTAGAGGACGCTGTTTAC 706
Qy 220 AsnAlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyxLeuGluGlnGluIle 239
Db 707 AATGTAATCAGGAGAAATATGCTATGAAAGAGACAGTGCATTTTGGAAACAGGAATA 766
Qy 240 LysGlyGluMetLysLeuLeuAsnAsnIleThrAsnAspLeuArgLeuLysAspTrpGlu 259
Db 767 AAAGAGAGAAATGAAAGTACTGAATAACATCACTAATGATCTCAGACTGAAAGATTGGAA 826
Qy 260 HisSerGlnThrLeuLysAsnIleThrLeuLeuGlnGly 272
Db 827 CATTTCTCAGACCTTGAGAAATATCACTTTAATTAATCAAGGT 865
RESULT 10
```

```

US-08-948-222-1
; Sequence 1, Application US/08948222
; Patent No. 5863798
; GENERAL INFORMATION:
; APPLICANT: Lysko, Paul G.
; APPLICANT: Elshourbagy, Nabil A.
; APPLICANT: Brawner, Mary E.
; TITLE OF INVENTION: Attachment Enhanced 293 Cells
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham - Corporate Patents
; ADDRESSEE: U.S.
; STREET: Mailcode - UW2220, 709 Swedeland Road
; CITY: King of Prussia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19406-5090
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/948,222
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/453,117
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Jervis, Herbert H.
; REGISTRATION NUMBER: 31,171
; REFERENCE/DOCKET NUMBER: SBC-P50338
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 270-5019
; TELEFAX: (610) 270-5090
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2028 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: No. 5863798 Relevant
; MOLECULE TYPE: cDNA to mRNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 47..1402
US-08-948-222-1
Alignment Scores:
Pred. No.: 1,5e-113 Length: 2028
Score: 1042.50 Matches: 197
Percent Similarity: 87.2% Conservativeness: 41
Best Local Similarity: 72.2% Mismatches: 34
Query Match: 50.6% Indels: 1
DB: 2 Gaps: 1

US-10-618-570-2 (1-400) x US-08-948-222-1 (1-2028)
Qy 1 MetAlaGlnTrpAspPheProAspGlnGlnGluAspThrAspSerCysThrGluSer 20
Db 47 ATGGAGCAGTGGGATCACTTTCAACAATCAACAGGAGGACACTGATAGCTGCTCCGAATCT 106
Qy 21 ValLysPheAspAlaAatGSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
Db 107 GTGAAATTTGATGCTCGCTCAATGACAGCTTTGCTTCCTCGAATCTTAAACACAGCCCT 166
Qy 41 ThrLeuGlnGluArgMetLysSerTyxLysThrAlaLeuLeuThrLeuTyxLeuIleVal 60
Db 167 TCCCTTCAAGAGAACTGAGTCTTCAAAAGCTGCACCTGATTTGCCCTTTACTCTCTCGTG 226
Qy 61 PheValValLeuValProIleGlyIleValAlaAlaGlnLeuLeuLeuTyxTrpGluThr 80
Db 227 TTTCAGTCTTCATCCCTCTCATTTGGAATAGTGGCAGCTCAACTCTGAAAGTGGAAACG 286
```


QY 81 LysAsnGlyThrValGlySerValAsnAla---AspIleSerProSerProGluGlyLys 99
 DB 287 AAGAAATTCCTCAGTTAGTTCACTAATCAATGATATACTCAAGTCTCACGGGAAA 346
 QY 100 GlyAsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsn 119
 DB 347 GGAATGACAGGAGAGAGAAATGATTTCAAGAGTCTTTATGGAACACATGAGCAAC 406
 QY 120 MetGluSerArgIleGlnTyrLeuSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsn 139
 DB 407 ATGAGAGAGAAATCCAGCATATTTAGACATGATGGAAGCCAACTCATGGACACAGAGCAT 466
 QY 140 PheGlnAsnPheSerIleThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsn 159
 DB 467 TTCACAAATTTTCAGCATGACAACTGATCAAAAGATTTAATGACATTTCTCGAGCTAAGT 526
 QY 160 SerLeuLeuSerSerIleGlnGluHisGluAsnIleIleGlyAspIleSerLysSerLeu 179
 DB 527 ACCTTGTTCTCAGTCAGGAGCATGGGAATGCAATAGATGAATCTCCAGTCCCTTA 586
 QY 180 ValGlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArg 199
 DB 587 ATAAGTTTGAATACCATGCTGTTGATTTGCGAGCTCAACATAGAAAATCTCAATGGCAA 646
 QY 200 ValGlnGluAsnAlaPheLysGlnGluMetArgLysLeuGluGluArgIleTyr 219
 DB 647 ATCAAGAGAGATACCTTCAACACAGAGGAAATCAGTAAATTAGAGGCGGTGTTTAC 706
 QY 220 AsnAlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyrLeuGluGlnGluIle 239
 DB 707 AATGTATCAGCAGAAATTTATGCTATGAAAGAGAACAAAGTGCAATTTGGAACAGGAATA 766
 QY 240 LysGlyGluMetLysLeuAsnAsnIleThrLeuLeuGlnGly 272
 DB 827 CATCTCAGACCTTGAGAAATATCATCTTAATTCAGT 865

RESULT 11

US-08-973-145-1
 ; Sequence 1, Application US/08973145
 ; Patent No. 5919636
 ; GENERAL INFORMATION:
 ; APPLICANT: Lysko, Paul G.
 ; APPLICANT: Elshourbagy, Nabil A.
 ; APPLICANT: Brawer, Mary E.
 ; TITLE OF INVENTION: Attachment Enhanced 293 Cells
 ; NUMBER OF SEQUENCES: 4
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: SmithKline Beecham - Corporate Patents U.S.
 ; STREET: Mailcode - UW2220, 709 Swedeland Road
 ; CITY: King of Prussia
 ; STATE: Pennsylvania
 ; COUNTRY: U.S.A.
 ; ZIP: 19406-5090
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/973,145
 ; FILING DATE: 26-NOV-1997
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Elizabeth J. Hecht
 ; REGISTRATION NUMBER: P-41, 824
 ; REFERENCE/DOCKET NUMBER: P50338
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (610) 270-5009

; TELFAX: (610) 270-5090
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2028 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: not relevant
 ; MOLECULE TYPE: cDNA to mRNA
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 47..1402
 ; US-08-973-145-1

Alignment Scores:
 Pred. No.: 1.5e-113 Length: 2028
 Score: 1042.50 Matches: 197
 Percent Similarity: 87.2% Conservative: 41
 Best Local Similarity: 72.2% Mismatches: 34
 Query Match: 50.6% Indels: 1
 DB: 2 Gaps: 1

US-10-618-570-2 (1-400) x US-08-973-145-1 (1-2028)

QY 1 MetAlaGlnTTPAspAspPheProAspGlnGlnGluAspThrAspSerCysThrGluSer 20
 DB 47 ATGAGAGAGTGGGATCACTTTCAATCAACAGAGGAGACACTGATAGTCTCCGAATCT 106
 QY 21 ValLysPheAspAlaArgSerValThrAlaLeuLeuProProHisProLysAsnGlyPro 40
 DB 107 GTGAATTTGATGCTCGCTCAATGACAGCTTGTCTCCGATCTTAAACACAGCCCT 166
 QY 41 ThrLeuGlnGluArgMetLysSerTyrLysThrAlaLeuIleThrLeuTyrLeuIleVal 60
 DB 167 TCCCTTCAAGAGAACTGAAGTCTTCAAGCTGCACCTGATTTGCCCTTTACCTCTCGTG 226
 QY 61 PheValValLeuValProIleIleGlyIleValAlaAlaGlnLeuLeuLysTyrGluThr 80
 DB 227 TTTGCAGTTCTCATCTCCCTCTCATTTGGAATAGTGGCAGCTCAACTCTGGAAGTGGAAACG 286
 QY 81 LysAsnGlyThrValGlySerValAsnAla---AspIleSerProSerProGluGlyLys 99
 DB 287 AAGAAATTCCTCAGTTAGTTCACTAATGCAATGATATACTCAAGTCTCACGGGAAA 346
 QY 100 GlyAsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsn 119
 DB 347 GGAATGACAGGAGAGAAATCAGATTTCAAGAGTCTTTATGGAACACATGAGCAAC 406
 QY 120 MetGluSerArgIleGlnTyrLeuSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsn 139
 DB 407 ATGAGAGAGAGATCCAGCATATTTTAGACATGGAAGCCAACTCATGGACACAGAGCAT 466
 QY 140 PheGlnAsnPheSerIleThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsn 159
 DB 467 TTCACAAATTTTCAGCATGACAACTGATCAAAAGATTTAATGACATTTCTCGAGCTAAGT 526
 QY 160 SerLeuLeuSerSerIleGlnGluHisGluAsnIleIleGlyAspIleSerLysSerLeu 179
 DB 527 ACCTTGTTCTCAGTCAGGAGCATGGGAATGCAATAGATGAATCTCCAGTCCCTTA 586
 QY 180 ValGlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArg 199
 DB 587 ATAAGTTTGAATACCATGCTGTTGATTTGCGAGCTCAACATAGAAAATCTCAATGGCAA 646
 QY 200 ValGlnGluAsnAlaPheLysGlnGluMetArgLysLeuGluGluArgIleTyr 219
 DB 647 ATCAAGAGAGATACCTTCAACACAGAGGAAATCAGTAAATTAGAGGCGGTGTTTAC 706
 QY 220 AsnAlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyrLeuGluGlnGluIle 239
 DB 707 AATGTATCAGCAGAAATTTATGCTATGAAAGAGAACAAAGTGCAATTTGGAACAGGAATA 766
 QY 240 LysGlyGluMetLysLeuAsnAsnIleThrLeuLeuGlnGly 259

Db 767 AAGGAGAGTGAAGTACTGAATAACATCACTAATGATCTCAGACTGAAGATTGGGAA 826
PCT-US96-08081-1
; Sequence 1, Application PC/TUS9608081
; GENERAL INFORMATION:
; APPLICANT: Smithline Beecham Corporation
; TITLE OF INVENTION: Attachment Enhanced 293 Cells
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Smithline Beecham - Corporate Patents U.S.
; STREET: Mailcode - U22220, 709 Swedeland Road
; CITY: King of Prussia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19406-5090
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/08081
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Jervis, Herbert H.
; REGISTRATION NUMBER: 31,171
; REFERENCE/DOCKET NUMBER: P50338
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 270-5019
; TELEFAX: (610) 270-5090
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2028 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: not relevant
; MOLECULE TYPE: cDNA to mRNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 47..1402
PCT-US96-08081-1
Alignment Scores:
Pred. No.: 1,5e-113 Length: 2028
Score: 1042.50 Matches: 197
Percent Similarity: 87.2% Conservative: 41
Best Local Similarity: 72.2% Mismatches: 34
Query Match: 50.6% Indels: 1
DB: 6 Gaps: 1
US-10-618-570-2 (1-400) x PCT-US96-08081-1 (1-2028)
Qy 1 MetAlaGlnTrpAspPheProAspGlnGlnGluAspThrAspSerCysThrGluSer 20
Db 47 ATGGAGCAGTGGGATCACTTTTCACAAATCAACAGGAGGACACTGATGCTCGGAATCT 106
Qy 21 VallysPheAspAlaArgSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
Db 107 GTGAAATTTGATGCTCGCTCAATGACAGCTTTGCTTCGGAATCTTAAACACAGCCCT 166
Qy 41 ThrLeuGlnGluArgMetLysSerTrpLysThrAlaLeuLeuThrLeuLeuVal 60
Db 167 TCCCTTCAAGAGAACTCAAGTCTTCAAGAGCTGCACTGATTCCTTACCTCTCTG 226
Qy 61 PheValValLeuValProIleGlyIleValAlaAlaGlnLeuLeuLysTrpGluThr 80
Db 227 TTTCAGTCTTCATCCCTCTCATTTGGAATAGTGGCAGCTCACTTCACTTGAAGTGGGAAACG 286

Qy 81 LysAsnCysThrValGlySerValAsnAla---AspIleSerProSerProGluGlyLys 99
Db 287 AAGAATTGCTCAGTTAGTTCAACTAATGCAATGATATNACTCAAGTCTCACGGGAAA 346
Qy 100 GlyAsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsn 119
Db 347 GGAATGCACAGCGAAGAGGAAATGAGATTTCAAGAAGTCTTTATGGAACACATGAGCAAC 406
Qy 120 MetGluSerArgIleGlnTrpLeuSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsn 139
Db 407 ATGGAGAAGAGATCCAGCATATTTTAGACATGGAAGCCACCTCATGGACACAGAGCAT 466
Qy 140 PheGlnAsnPheSerIleThrThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsn 159
Db 467 TTCCAAATTTTACGATGACACACTGATCAAAAGATTTAATGACATTTCTTCGACGTAAGT 526
Qy 160 SerLeuLeuSerSerIleGlnGluHisGluAsnIleIleGlyAspIleSerLysSerLeu 179
Db 527 ACCTTGTCTTCAGTCCAGGACATGGGAATGCAATAGATGAATCTCCAAGTCCCTTA 586
Qy 180 ValGlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArg 199
Db 587 ATAAAGTTTGAATACCACTTGTGATTTCAGCTCAACATAGAAAATCTGAATGGCAAA 646
Qy 200 ValGlnGluAsnAlaPheLysGlnGlnGluMetArgLysLeuGluArgIleTyr 219
Db 647 ATCCAGAGAAATACCTTCAACAAAGAGAAATCAGTAATATTAGAGGACGCTGTTTAC 706
Qy 220 AsnAlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyrLeuGluGlnGluLe 239
Db 707 AATGATATCAGCAGAAATTTATGCTATGAAAGAGACACAGTGCATTTTGGACAGGAAATA 766
Qy 240 LysGlyGluMetLysLeuLeuAsnAsnIleThrAsnAspLeuArgLeuLysAspTrpGlu 259
Db 767 AAGGAGAGTGAAGTACTGAATAACATCACTAATGATCTCAGACTGAAAGATTGGGAA 826
Qy 260 HisSerGlnThrLeuLysAsnIleThrLeuLeuGlnGly 272
Db 827 CATCTCAGACCTTGAAGAAATATCACTTTAATTCAGGT 865
RESULT 13
US-08-154-365-1
; Sequence 1, Application US/08154365
; Patent No. 5624904
; GENERAL INFORMATION:
; APPLICANT: Dunne, Dana W.
; APPLICANT: Resnick, David
; APPLICANT: Kreiger, Monty
; APPLICANT: Joiner, Keith A.
; TITLE OF INVENTION: Method for Treating Gram-Positive
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 1100 Peachtree Street, Suite 2800
; CITY: Atlanta
; STATE: Ga
; COUNTRY: USA
; ZIP: 30309-4530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/154,365
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: MIT6392

TELECOMMUNICATION INFORMATION:

TELEPHONE: (404)-815-6508

TELEFAX: (404)-815-6555

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 2037 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: homo sapien

IMMEDIATE SOURCE:

LIBRARY: THP-1

PUBLICATION INFORMATION:

AUTHORS: Ashkenas, et al.

JOURNAL: J. Lipid Res.

VOLUME: 34

PAGES: 983-1000

DATE: 1993

RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 2037

US-08-154-365-1

Alignment Scores:

| | | | |
|------------------------|-----------|---------------|------|
| Pred. No.: | 3.44e-113 | Length: | 2037 |
| Score: | 1039.50 | Matches: | 196 |
| Percent Similarity: | 87.2% | Conservative: | 42 |
| Best Local Similarity: | 71.8% | Mismatches: | 34 |
| Query Match: | 50.5% | Indels: | 1 |
| DB: | 2 | Gaps: | 1 |

US-10-618-570-2 (1-400) x US-08-154-365-1 (1-2037)

```

Qy 1 MetAlaGlnTTPApAspPheProAspGlnGluAspThrAspSerCysThrGluSer 20
Db 47 ATGGAGCAGTGGGATCACTTTTCAATCAATCAACAGGAGGACACTGATAGCTGCTCGGAATCT 106
Qy 21 ValLysPheAspAlaArgSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
Db 107 GTGAATTTGATGCTCGCTCAATGACAGCTTGTCTTCCGATCTCTAATAAAGAGCCT 166
Qy 41 ThrLeuGlnGluArgMetLysSerTyrLysThrAlaLeuLeuThrLeuLysLeuVal 60
Db 167 TCCCTTCAAGAGAACTCAAGTCTTCAAGCTGCACTGATTCCTTACCTTCTCTGCTG 226
Qy 61 PheValValLeuValProIleGlyTleValAlaGlnLeuLeuLysTTPGluThr 80
Db 227 TTTCAGGTTCATCCCTCTCATTTGGAATAGTGCAGCTCAACTCCTGAAGTGGGAAACG 286
Qy 81 LysAsnCysThrValGlySerValAsnAla--AspIleSerProSerProGluGlyLys 99
Db 287 AAGATTGCTCAGTAGTTCAGTAATGCAATATGATATCAAGTCTCAGCGGAAA 346
Qy 100 GlyAsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsn 119
Db 347 GGAATGACGCGGAGAGAAATGAGATTTCAGAAAGTCTTTATGCAACACATCAGCAAC 406
Qy 120 MetGluSerArgIleGlnTyrLeuSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsn 139
Db 407 ATGGAGAGAGAAATCCAGCATATTTTAGCATGGAAGCCAACTCATGGACACAGAGCAT 466
Qy 140 PheGlnAsnPheSerIleThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsn 159
Db 467 TTCGAAATTTTCAGATGACAACTGATCAAGATTTAATGACATTTCTCGCATAGT 526
Qy 160 SerLeuLeuSerSerIleGlnGluHisGluAsnIleIleGlyAspIleSerLysSerLeu 179
Db 527 ACCTTGTTTCTCAGTCCAGGACATGGGAATGCAATAGATGAATCAATCTCAAGTCTCTTA 586
Qy 180 ValGlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArg 199

```

```

Db 587 ATAAGTTTGAATACCACTTGCCTTGTGATTTCAGCTCAACATAGAAATCTGAATGGCAA 646
Qy 200 ValGlnGluAsnAlaPheLysGlnGlnGluMetArgLysLeuGluGluArgIleTyr 219
Db 647 ATCCAAGAGAATACCTTCAAAACAAGAGAAATCAGTAATTAGAGGACGCTGTTTAC 706
Qy 220 AsnAlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyrLeuGluGluGluIle 239
Db 707 AATGTATCAGCAGAAATATTGGCTATGAAAGAAACAAGTGCATTTTGGAAACAGGAATA 766
Qy 240 LysGlyGluMetLysLeuLeuAsnAsnIleThrAsnAspLeuArgLeuLysAspTTPGlu 259
Db 767 AAGAGAGAGTGAAGTACTGATTAACATCACTAATGATCTGACACTGGAAGATTGGGA 826
Qy 260 HisSerGlnThrLeuLysAsnIleThrLeuLeuGlnGly 272
Db 827 CATTCAGACCTTGAGAAATATCATTTAATCAAGT 865

```

RESULT 14

US-08-554-586-1

Sequence 1, Application US/08554586

Patent No. 5767379

GENERAL INFORMATION:

APPLICANT: BASZCZYNSKI, Chris

APPLICANT: HOOD, Elizabeth

APPLICANT: MADDOCK, Sheila

APPLICANT: MEYER, Terry Euclyaire

APPLICANT: REGISTER III, James C.

APPLICANT: WITCHER, Derrick

APPLICANT: HOWARD, John A.

TITLE OF INVENTION: COMMERCIAL PRODUCTION OF AVIDIN IN

NUMBER OF SEQUENCES: 3

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Foley & Lardner

STREET: 3000 K Street, N.W., Suite 500

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20007-5109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/554,586

FILING DATE: 06-NOV-1995

CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:

NAME: BENT, Stephen A.

REGISTRATION NUMBER: 29,768

REFERENCE/DOCKET NUMBER: 33229/352/PIHI

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202)672-5300

TELEFAX: (202)672-5399

TELEX: 904136

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 484 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

US-08-554-586-1

Alignment Scores:

| | | | |
|------------------------|----------|---------------|-----|
| Pred. No.: | 2.29e-72 | Length: | 484 |
| Score: | 688.00 | Matches: | 130 |
| Percent Similarity: | 98.5% | Conservative: | 1 |
| Best Local Similarity: | 97.7% | Mismatches: | 2 |
| Query Match: | 33.4% | Indels: | 0 |
| DB: | 2 | Gaps: | 0 |

US-10-618-570-2 (1-400) x US-08-554-586-1 (1-484)

```
QY 268 ThrLeuLeuGlnGlyAlaArgLysCysSerLeuThrGlyLysThrAsnAspLeuGly 287
Db 69 TCCCTCGCCAGCGCGCCAGGAAGTCTCCCTCACCGCAAGTGGACCAATGACTCGGC 128
QY 288 SerAsnMetThrIleGlyAlaValAsnSerArgGlyGluPheThrGlyThrTyrlleThr 307
Db 129 TCCACATGACCATCGCGCGCGTGAACCTCAGGGGGGAGTTCCCGGCACCTACATCACC 188
QY 308 AlaValThrAlaThrSerAsnGluLeuLysGluSerProLeuHisGlyThrGlnAsnThr 327
Db 189 GCGGTGACCGGCACCTCCACAGCAGATCAAGAGTCCCCCTCCACGGTACCCAGACACC 248
QY 328 IleAsnLysArgThrGlnProThrPheGlyPheThrValAsnTrpLysPheSerGluSer 347
Db 249 ATCAACAAGAGGACCCAGCCACCTTCGGCTTCAACCGTGAACCTGGAAAGTTCTCCGAGTCC 308
QY 348 ThrThrValPheThrGlyGlnCysPheIleAspArgAsnGlyLysGluValLeuLysThr 367
Db 309 ACCACCGTGTTCACCGCGCAGTCTTCATCCAGCCACCGCAAGGAGTCTCAAGACC 368
QY 368 MetTrpLeuLeuArgSerSerValAsnAspIleGlyAspAspTrpLysAlaThrArgVal 387
Db 369 ATGTGGCTCTGAGGAGCTCCGTGAATGACATCGCGACGACTCGAAGGCCACCCGCGTG 428
QY 388 GlyIleAsnIlePheThrArgLeuArgThrGlnLysGlu 400
Db 429 GGCATCAACATCTTCACCCGCTCCGACCCAGAGGAG 467
```

RESULT 15

```
US-08-831-399-3
; Sequence 3, Application US/08831399
; Patent No. 6312916
; GENERAL INFORMATION:
; APPLICANT: Kopeckzi, Erhard; Muller, Rainer;
; APPLICANT: Engh, Richard; Schmitt, Urban; Deger, Arno; Brandstetter, Hans
; TITLE OF INVENTION: Recombinant Inactive Core
; TITLE OF INVENTION: Streptavidin Mutants
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
; COMPUTER: IBM PS/2 PC-DOS
; OPERATING SYSTEM:
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/831.399
; FILING DATE: 1-April-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 196 13 053.0
; FILING DATE: 1-April-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 196 37 718.8
; FILING DATE: 16-September-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 6312916man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: HUBR 1105
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 604 base pairs
; TYPE: nucleic acid
```

```
; STRANDEDNESS: double
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 44..499
; OTHER INFORMATION: Positions 44..115 correspond to sig
; OTHER INFORMATION: peptide and 116..499 to mat peptide
US-08-831-399-3
```

```
Alignment Scores:
Pred. No.: 5,07e-71 Length: 604
Score: 678.00 Matches: 128
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 32.9% Indels: 0
DB: 3 Gaps: 0
```

US-10-618-570-2 (1-400) x US-08-831-399-3 (1-604)

```
QY 273 AlaArgLysCysSerLeuThrGlyLysThrAsnAspLeuGlySerAsnMetThrIle 292
Db 116 GCAGAAAGTGTCTGCTGACTGGGAAATGGACCAACGATCTGGGCTCCAACATGACCATC 175
QY 293 GlyAlaValAsnSerArgGlyGluPheThrGlyThrTyrlleThrAlaValThrAlaThr 312
Db 176 GGGGCTGTGAACAGCAGAGGTGAATTCACAGGCACCTACATCACAGCGTAACAGCCACA 235
QY 313 SerAsnGluLeuLysGluSerProLeuHisGlyThrGlnAsnThrIleAsnLysArgThr 332
Db 236 TCAATGAGATCAAAAGAGTCAACCACTGCATGGGACACAAAACACCATCAACAGAGGACC 295
QY 333 GlnProThrPheGlyPheThrValAsnTrpLysPheSerGluSerThrValPheThr 352
Db 296 CAGCCACCTTTGGCTTTCACCGCAATGGAAAGTTTTCAGAGATCCACCATCTTTCACG 355
QY 353 GlyGlnCysPheIleAspArgAsnGlyLysGluValLeuLysThrMetTrpLeuLeuArg 372
Db 356 GGCAGTGTCTCATAGACAGGAATGGGAAGAGGTCTCTGAAGACCATGTGGCTGTGGGG 415
QY 373 SerSerValAsnAspIleGlyAspAspTrpLysAlaThrArgValGlyIleAsnIlePhe 392
Db 416 TCAAGTGTTAATGACATTTGGTGATGACTGGAAAGCTACCAAGGCTACCAAGGCTCGGCATCAACATCTTC 475
QY 393 ThrArgLeuArgThrGlnLysGlu 400
Db 476 ACTGCGCTCGGCACACAGAAAGGAG 499
```

Search completed: February 27, 2006, 12:42:35
Job time : 239 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 27, 2006, 08:22:53 ; Search time 24993 Seconds
(without alignments)
11774.439 Million cell updates/sec

Title: US-10-618-570-1
Perfect score: 5177
Sequence: 1 ttgaagaccaccacgta.....agcgggggttttcattgg 5177

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

1: gb_ba.*
2: gb_in.*
3: gb_env.*
4: gb_on.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pr.*
9: gb_ro.*
10: gb_sts.*
11: gb_sy.*
12: gb_un.*
13: gb_vi.*
14: gb_htg.*
15: gb_pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|-------------|
| 1 | 5177 | 100.0 | 5177 | 6 | BD136828 |
| 2 | 5177 | 100.0 | 5177 | 6 | AX018983 |
| 3 | 2078.8 | 40.2 | 7699 | 6 | A80474 |
| 4 | 2078.8 | 40.2 | 7699 | 6 | BD136838 |
| 5 | 2078.8 | 40.2 | 7699 | 6 | AX381629 |
| 6 | 1972.6 | 38.1 | 7436 | 6 | C0879093 |
| 7 | 1916.6 | 37.0 | 7546 | 6 | BD138566 |
| 8 | 1644.6 | 31.8 | 7165 | 6 | BD138564 |
| 9 | 1629.4 | 31.5 | 5594 | 6 | AR214682 |
| 10 | 1588.6 | 30.7 | 5130 | 6 | AR634978 |
| 11 | 1588.6 | 30.7 | 5130 | 6 | AX359935 |
| 12 | 1588.6 | 30.7 | 5130 | 6 | AX382149 |
| 13 | 1487.6 | 28.7 | 5464 | 11 | SYNMMPLN1 |
| 14 | 1486 | 28.7 | 5874 | 6 | BD138565 |
| 15 | 1486 | 28.7 | 5874 | 11 | SYNMMPLN4 |
| 16 | 1486 | 28.7 | 6141 | 6 | BD138562 |
| 17 | 1486 | 28.7 | 6365 | 6 | I70974 |
| 18 | 1486 | 28.7 | 6506 | 13 | U00220 |

| | | | | | |
|----|--------|------|------|----|----------|
| 19 | 1486 | 28.7 | 6522 | 6 | BD138563 |
| 20 | 1486 | 28.7 | 6620 | 6 | AR302094 |
| 21 | 1463 | 28.3 | 6501 | 11 | AF132211 |
| 22 | 1455.2 | 28.1 | 5865 | 6 | AR164477 |
| 23 | 1452 | 28.0 | 6072 | 11 | D88622 |
| 24 | 1448.4 | 28.0 | 8591 | 6 | I58322 |
| 25 | 1448.4 | 28.0 | 8591 | 6 | I58323 |
| 26 | 1448.4 | 28.0 | 8591 | 6 | I60508 |
| 27 | 1448.4 | 28.0 | 8591 | 6 | I60509 |
| 28 | 1448.4 | 28.0 | 8591 | 6 | I77052 |
| 29 | 1448.4 | 28.0 | 8591 | 6 | I77053 |
| 30 | 1448.4 | 28.0 | 8591 | 6 | I87173 |
| 31 | 1448.4 | 28.0 | 8591 | 6 | I87174 |
| 32 | 1448 | 28.0 | 6094 | 11 | AF113968 |
| 33 | 1447.4 | 28.0 | 6248 | 11 | AB041928 |
| 34 | 1444.6 | 27.9 | 6277 | 11 | AB086385 |
| 35 | 1437.2 | 27.8 | 9609 | 11 | AP396261 |
| 36 | 1434 | 27.7 | 6700 | 6 | AR585288 |
| 37 | 1434 | 27.7 | 8518 | 6 | AR321727 |
| 38 | 1434 | 27.7 | 8518 | 6 | AR585290 |
| 39 | 1434 | 27.7 | 8518 | 6 | AR656176 |
| 40 | 1429.2 | 27.6 | 6700 | 6 | AR321726 |
| 41 | 1429.2 | 27.6 | 6700 | 6 | AR656175 |
| 42 | 1425 | 27.5 | 5689 | 6 | BD075836 |
| 43 | 1425 | 27.5 | 5689 | 6 | BD081937 |
| 44 | 1425 | 27.5 | 5689 | 6 | BD269812 |
| 45 | 1425 | 27.5 | 5689 | 6 | AR593118 |

ALIGNMENTS

RESULT 1
BD136828
LOCUS BD136828 5177 bp DNA linear PAT 18-SEP-2002
DEFINITION Biotin-binding receptor molecule.
ACCESSION BD136828
VERSION BD136828.1 GI:23231773
KEYWORDS JP 2002504328-A/1.
SOURCE unidentified
ORGANISM unclassified.

REFERENCE 1 (bases 1 to 5177)
AUTHORS Hertzuala,S.Y., Kulomaa,M., Lehtolainen,P., Marjomaki,V. and Airenne,K.
TITLE Biotin-binding receptor molecule
JOURNAL Patent: JP 2002504328-A 1 12-FEB-2002;
ARC THERAPEUTICS LTD
COMMENT OS Unidentified
PN JP 2002504328-A/1
PD 12-FEB-2002
PF 23-FEB-1999 JP 2000532517
PR 23-FEB-1998 GB 9803757.5,24-JUN-1998 GB 9813653 4 PI
SPPPO YLA HERTZUALA,MARKKU KULOMAA,PAULIINA LEHTOLAINEN,VARPU PI
MARJOMAKI,
PI KARI AIRENNE
PC C12N15/09,A61K48/00,A61P43/00,C07K14/705//A61K38/00,C12N15/00,
PC A61K37/02
CC Strandedness: Single;
CC Topology: linear;
CC Biotin-binding receptor molecule
FH Key Location/Qualifiers
FT CDS 1071..2270.
Location/Qualifiers
source
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN

Query Match 100.0%; Score 5177; DB 6; Length 5177;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTTGAAAGACCCCAACCCGTAGGTGGCAAGCTAGCTTAAGTAAACGCCACTTTTGCAGGCAT 60
Db |||||
1 TTTGAAAGACCCCAACCCGTAGGTGGCAAGCTAGCTTAAGTAAACGCCACTTTTGCAGGCAT 60
Qy 61 GGAATAATACATAATCTAGAAATAGAAAGTTTCAGATCAAGGTTCAGGAACAAGAAACAGC 120
Db |||||
61 GGAATAATACATAATCTAGAAATAGAAAGTTTCAGATCAAGGTTCAGGAACAAGAAACAGC 120
Qy 121 TGAATACCAACACAGATATCTGTGGTAAGCGGTTCTGCCCGGCTCAGGCCCAAGAAACA 180
Db |||||
121 TGAATACCAACACAGATATCTGTGGTAAGCGGTTCTGCCCGGCTCAGGCCCAAGAAACA 180
Qy 181 GATGAGACAGCTGAGTGTATGGGCCCAACACAGATATCTGTGGTAAGCGGTTCTGCCCGG 240
Db |||||
181 GATGAGACAGCTGAGTGTATGGGCCCAACACAGATATCTGTGGTAAGCGGTTCTGCCCGG 240
Qy 241 CTCGGGCCCAAGAAACAGATGTCTCCAGATGCGGTCCAGCCCTCAGCAGTTTCTAGTGA 300
Db |||||
241 CTCGGGCCCAAGAAACAGATGTCTCCAGATGCGGTCCAGCCCTCAGCAGTTTCTAGTGA 300
Qy 301 TCATCAGATGTTTCCAGGGTCCCAAGGACCTGAAATGACCCCTGTACCTTATTTGAAC 360
Db |||||
301 TCATCAGATGTTTCCAGGGTCCCAAGGACCTGAAATGACCCCTGTACCTTATTTGAAC 360
Qy 361 TAAACAATCAGTTCGCTTCTCGCTTCTGTTTCGCGCTTCCGCTCTCCGAGCTCAATAAA 420
Db |||||
361 TAAACAATCAGTTCGCTTCTCGCTTCTGTTTCGCGCTTCCGCTCTCCGAGCTCAATAAA 420
Qy 421 AGAGCCCAACACCCCTCACTCGGCGGCCAGTCTTCCGATAGACTGCGTCCGCCGGTAC 480
Db |||||
421 AGAGCCCAACACCCCTCACTCGGCGGCCAGTCTTCCGATAGACTGCGTCCGCCGGTAC 480
Qy 481 CCGTATTTCCCAATAAAGCTCTTGGTTCGATTCGAAATCGTGGTCTCGCTGTTCTCTTG 540
Db |||||
481 CCGTATTTCCCAATAAAGCTCTTGGTTCGATTCGAAATCGTGGTCTCGCTGTTCTCTTG 540
Qy 541 GGAGGGTCTCTCTGAGTGTAGTACTACCAACAGCGGGGTCTTTTCAATTTGGGGGCTCGT 600
Db |||||
541 GGAGGGTCTCTCTGAGTGTAGTACTACCAACAGCGGGGTCTTTTCAATTTGGGGGCTCGT 600
Qy 601 CCGGGATTTGGAGACCCCTGCCCAGGACCAACCGACCCACCGAGGAGTAAAGCTGGCC 660
Db |||||
601 CCGGGATTTGGAGACCCCTGCCCAGGACCAACCGACCCACCGAGGAGTAAAGCTGGCC 660
Qy 661 AGCAACTTATCTGTCTGTCGGAATGCTAGTGTCTATGTTGATGTTATGCGCCTGCG 720
Db |||||
661 AGCAACTTATCTGTCTGTCGGAATGCTAGTGTCTATGTTGATGTTATGCGCCTGCG 720
Qy 721 TCTGTACTAGTTAGCTTAAGCTCTGTA TCTGGCGACCCGCTGCGAACTGACGAGTT 780
Db |||||
721 TCTGTACTAGTTAGCTTAAGCTCTGTA TCTGGCGACCCGCTGCGAACTGACGAGTT 780
Qy 781 CTGAACACCCCGCCGCAACCCCTGGGAGACGTTCCAGGGACTTTGGGGGCCGTTTGTGG 840
Db |||||
781 CTGAACACCCCGCCGCAACCCCTGGGAGACGTTCCAGGGACTTTGGGGGCCGTTTGTGG 840
Qy 841 CCCGACCTGAGGAAGGAGTGCATGTGGAATCCGACCCCGCTCAGGATATGTGGTCTGGT 900
Db |||||
841 CCCGACCTGAGGAAGGAGTGCATGTGGAATCCGACCCCGCTCAGGATATGTGGTCTGGT 900
Qy 901 AGGAGAGGAAACCTAAACAGTTTCCGCGCTCCGCTCTGAAATTTTTCGTTTCGGTTTGGAA 960
Db |||||
901 AGGAGAGGAAACCTAAACAGTTTCCGCGCTCCGCTCTGAAATTTTTCGTTTCGGTTTGGAA 960
Qy 961 CCGAAGCGCGGCTTGTCTGTCTGCAGCCCAAGCTTGGGGTGCAGGTCCGACTTAGAGGA 1020
Db |||||
961 CCGAAGCGCGGCTTGTCTGTCTGCAGCCCAAGCTTGGGGTGCAGGTCCGACTTAGAGGA 1020
Qy 1021 TCAATTCGGCACGAGTAAATCGGTGCTGCCGTCTTTTAGGACATATGAAGTATGCAAGT 1080
Db |||||
1021 TCAATTCGGCACGAGTAAATCGGTGCTGCCGTCTTTTAGGACATATGAAGTATGCAAGT 1080

Qy 1081 GGGATGATCTTCTCGATCAGCAAGAGGACACTGACAGCTGTACAGAGTCTGTGAAGTTGC 1140
Db |||||
1081 GGGATGATCTTCTCGATCAGCAAGAGGACACTGACAGCTGTACAGAGTCTGTGAAGTTGC 1140
Qy 1141 ATGCTCGCTCAGTGACAGCTTTGCTTCTCCCATCTCTAAATATGCCCCAACTCTTCAAG 1200
Db |||||
1141 ATGCTCGCTCAGTGACAGCTTTGCTTCTCCCATCTCTAAATATGCCCCAACTCTTCAAG 1200
Qy 1201 AGAGGATGAAGTCTTATAAACTGCACTGATCACTCCCTTTATCTCATTTGTTGTAGTTC 1260
Db |||||
1201 AGAGGATGAAGTCTTATAAACTGCACTGATCACTCCCTTTATCTCATTTGTTGTAGTTC 1260
Qy 1261 TCGTGCCCATCAATGGCATAGTGGCAGCTCAGCTCTCTGAAATGGGAAACGAAGAAATTCGA 1320
Db |||||
1261 TCGTGCCCATCAATGGCATAGTGGCAGCTCAGCTCTCTGAAATGGGAAACGAAGAAATTCGA 1320
Qy 1321 CGTTTGGCTCAGTTAATGCAAGATATATCTCAAAGTCCGGAAGGCAAGAAATGGCAGTG 1380
Db |||||
1321 CGTTTGGCTCAGTTAATGCAAGATATATCTCAAAGTCCGGAAGGCAAGAAATGGCAGTG 1380
Qy 1381 AAGATGAAATGAGATTTGAGAAAGCTGTGATGGAACGCGATGAGCAACATCGAAAGCAGAA 1440
Db |||||
1381 AAGATGAAATGAGATTTGAGAAAGCTGTGATGGAACGCGATGAGCAACATCGAAAGCAGAA 1440
Qy 1441 TCCAGTATCTTTCAGATAATGAAGCCAACTCTCTAGATGCTAAGAAATTTCCAAAATTTCA 1500
Db |||||
1441 TCCAGTATCTTTCAGATAATGAAGCCAACTCTCTAGATGCTAAGAAATTTCCAAAATTTCA 1500
Qy 1501 GCATAACAACCTGATCAAAAGTTAATGATGTTCTTTTCCAGCTTAAATCTCTTACTTTCT 1560
Db |||||
1501 GCATAACAACCTGATCAAAAGTTAATGATGTTCTTTTCCAGCTTAAATCTCTTACTTTCT 1560
Qy 1561 CCATCAGGACATGAGATATCATAGGGGATATCTCCAGTCTNTAGTAGGCTGGAACA 1620
Db |||||
1561 CCATCAGGACATGAGATATCATAGGGGATATCTCCAGTCTNTAGTAGGCTGGAACA 1620
Qy 1621 CCACAGTACTTGTATTTGCAAGTTCAGTATTTGAAACACATGAAATGGCAGAGTCAAGAGAA 1680
Db |||||
1621 CCACAGTACTTGTATTTGCAAGTTCAGTATTTGAAACACATGAAATGGCAGAGTCAAGAGAA 1680
Qy 1681 CATTAAACAACAAGAGGAGATCGTAAATTTAGAGGAGCGTATATACAATGCAATCAGCAG 1740
Db |||||
1681 CATTAAACAACAAGAGGAGATCGTAAATTTAGAGGAGCGTATATACAATGCAATCAGCAG 1740
Qy 1741 AAATTAAGTCTCTAGATGAAACAAAGTATATTTTGGAAACAGGAAATAAAGGGGAAATGA 1800
Db |||||
1741 AAATTAAGTCTCTAGATGAAACAAAGTATATTTTGGAAACAGGAAATAAAGGGGAAATGA 1800
Qy 1801 AACCTGTTGAATAATATATCACTAATGATCTGAGGCTGAAAGGATTTGGGAACTTCTCAGACAT 1860
Db |||||
1801 AACCTGTTGAATAATATATCACTAATGATCTGAGGCTGAAAGGATTTGGGAACTTCTCAGACAT 1860
Qy 1861 TGAATAATATCACTTTTACTCCAAAGTGGCAGAAAGTGTCTGCTGA CTGGGAAATGGACCA 1920
Db |||||
1861 TGAATAATATCACTTTTACTCCAAAGTGGCAGAAAGTGTCTGCTGA CTGGGAAATGGACCA 1920
Qy 1921 ACCATCTGGGCTCAACATGACCATCGGGCTGTGAACAGCAGAGGTAATTCACAGGCA 1980
Db |||||
1921 ACCATCTGGGCTCAACATGACCATCGGGCTGTGAACAGCAGAGGTAATTCACAGGCA 1980
Qy 1981 CCTACATCACAGCGTAAACAGCCACATCAATGAGATCAAAAGTCAACACTGATCGGGA 2040
Db |||||
1981 CCTACATCACAGCGTAAACAGCCACATCAANTGAGATCAAAAGTCAACACTGATCGGGA 2040
Qy 2041 CACAAAACCACTCAACAAAGAGGACCCAGCCCACTTTGGCTTCA CCGTCAATTTGGAAGT 2100
Db |||||
2041 CACAAAACCACTCAACAAAGAGGACCCAGCCCACTTTGGCTTCA CCGTCAATTTGGAAGT 2100
Qy 2101 TTTTCAAGTCCACACTGTCTTCA CCGGCCAGTGTCTTCA TAGACAGGAATGGGAGAGG 2160
Db |||||
2101 TTTTCAAGTCCACACTGTCTTCA CCGGCCAGTGTCTTCA TAGACAGGAATGGGAGAGG 2160
Qy 2161 TCCTGAAGACCATGTGGCTGCTCGGCTCAAGTGTGTTAATGACATTTGGTGTGATGACTGGAAAG 2220

Db 2161 TCCTGAAGACCATGTGGCTGCTGGCGTCAAGTGTTAATGACATTTGGTGATGACTGGAAAG 2220
Qy 2221 CTACAGGGTCGGCATCAACATCTTTCACTCGCCCTGGCGCACAGAGAGAGTGAAGTG 2280
Db 2221 CTACAGGGTCGGCATCAACATCTTTCACTCGCCCTGGCGCACAGAGAGAGTGAAGTG 2280
Qy 2281 ACCAAGGTCTCTCTGGAATCCAGGTGAAGGAGAGATAGAGGCGCTCTCTGACAAAATGG 2340
Db 2281 ACCAAGGTCTCTCTGGAATCCAGGTGAAGGAGAGATAGAGGCGCTCTCTGACAAAATGG 2340
Qy 2341 TATACCAAGCTTTCCAGGTCTAAATAGGTACTCCAGGTCTTAAAGGTGATCGGGGGATCT 2400
Db 2341 TATACCAAGCTTTCCAGGTCTAAATAGGTACTCCAGGTCTTAAAGGTGATCGGGGGATCT 2400
Qy 2401 CTGGTTTACCTGGAGTTCCAGGATTTCCAGGACCAATGGGGAAGACCGGGAAGCCAGGAC 2460
Db 2401 CTGGTTTACCTGGAGTTCCAGGATTTCCAGGACCAATGGGGAAGACCGGGAAGCCAGGAC 2460
Qy 2461 TTAATGGACAAAAGGCGCAGAGGAGAAAAGGGAGTGGGAAGCATGCAAAAGACAATCTTA 2520
Db 2461 TTAATGGACAAAAGGCGCAGAGGAGAAAAGGGAGTGGGAAGCATGCAAAAGACAATCTTA 2520
Qy 2521 ATACAGTCCGACTGGTGGGTCGAGCGCCCTCACGAGGCAGAGTGGAGATTTTTCACG 2580
Db 2521 ATACAGTCCGACTGGTGGGTCGAGCGCCCTCACGAGGCAGAGTGGAGATTTTTCACG 2580
Qy 2581 AAGGCCAGTGGGTAACGCTGTGACGACCGCTCGGAACCTCGTGGAGGACTGCTGCTCT 2640
Db 2581 AAGGCCAGTGGGTAACGCTGTGACGACCGCTCGGAACCTCGTGGAGGACTGCTGCTCT 2640
Qy 2641 GCAGGAGCTTGGGATCAAAAGGTGTTCAAAGTGTGCTAAAGCGAGCTTATTTTGGAAAAG 2700
Db 2641 GCAGGAGCTTGGGATCAAAAGGTGTTCAAAGTGTGCTAAAGCGAGCTTATTTTGGAAAAG 2700
Qy 2701 GTAAGGTCCAAATAGCTGTAATGCTGTAATTTTGTTCGGGAAGAGTATCCATTTGAAG 2760
Db 2701 GTAAGGTCCAAATAGCTGTAATGCTGTAATTTTGTTCGGGAAGAGTATCCATTTGAAG 2760
Qy 2761 AGTCAGAAATPAGACAGTGGGCTGTGAGAGCTGTTCGACAGCAAGAGATGCTGGGGTC 2820
Db 2761 AGTCAGAAATPAGACAGTGGGCTGTGAGAGCTGTTCGACAGCAAGAGATGCTGGGGTC 2820
Qy 2821 ACTTTGCACTTACATAATATGCTATTTTCAATTTTCAATTTTAAAGTGT 2880
Db 2821 ACTTTGCACTTACATAATATGCTATTTTCAATTTTCAATTTTAAAGTGT 2880
Qy 2881 ATTTTCTTCTTCTTCACTAAATCAGCTTAATTAATTTTAAAGAACTTAAGAAATTTT 2940
Db 2881 ATTTTCTTCTTCTTCACTAAATCAGCTTAATTAATTTTAAAGAACTTAAGAAATTTT 2940
Qy 2941 ATCCACAGAAAGGAATATTTTAAATACTACCTGGATAAATATAGCTTCAATTTT 3000
Db 2941 ATCCACAGAAAGGAATATTTTAAATACTACCTGGATAAATATAGCTTCAATTTT 3000
Qy 3001 GCTTCAATACAGAACCAATTTCACTTCTAGGTGTTTTTAAAGTGGCTCGTGGCGAATG 3060
Db 3001 GCTTCAATACAGAACCAATTTCACTTCTAGGTGTTTTTAAAGTGGCTCGTGGCGAATG 3060
Qy 3061 ATCCCTCAGGATATAGTATTTTCTGCTTTCGATAGGAGGGGGAATATGATGCTTATGC 3120
Db 3061 ATCCCTCAGGATATAGTATTTTCTGCTTTCGATAGGAGGGGGAATATGATGCTTATGC 3120
Qy 3121 AATACTCTTGTAGTCTTGAACATGTTAAGTGTAGCAATGCTTCTTACAGGGA 3180
Db 3121 AATACTCTTGTAGTCTTGAACATGTTAAGTGTAGCAATGCTTCTTACAGGGA 3180
Qy 3181 GAAAAGCAGCCGTCAGTCCGATTTGGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 3240
Db 3181 GAAAAGCAGCCGTCAGTCCGATTTGGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 3240
Qy 3241 AAGCAACAGACGGGTCTGACATGGAATGGACGAACCACTGAATTCGCGATTCAGAGAT 3300

Db 3241 AAGCAACAGACGGGTCTGACATGGATTGGACGAACCACTGAATTCGCGATTCAGAGAT 3300
Qy 3301 ATTGTATTTAAGTTCCTAGCTCGATACAGCAAAAGCCATTTAGCATTTACCAATTTGGT 3360
Db 3301 ATTGTATTTAAGTTCCTAGCTCGATACAGCAAAAGCCATTTAGCATTTACCAATTTGGT 3360
Qy 3361 GTGCACCTCCAGGTTTCAAGCTTCGCGCAAGCACTCAGGGCGCAAGGGCTGCTAAAGGAAG 3420
Db 3361 GTGCACCTCCAGGTTTCAAGCTTCGCGCAAGCACTCAGGGCGCAAGGGCTGCTAAAGGAAG 3420
Qy 3421 CGGAACACGTAGAAGCAGTCCGCGAGAAACGGTGTGACCCCGGATGAATGTCAAGCTAC 3480
Db 3421 CGGAACACGTAGAAGCAGTCCGCGAGAAACGGTGTGACCCCGGATGAATGTCAAGCTAC 3480
Qy 3481 TGGGCTATCTGGACAAAGGAAAACGCAAGCGCAAGAGAAAGCAAGTGTGAGTGG 3540
Db 3481 TGGGCTATCTGGACAAAGGAAAACGCAAGCGCAAGAGAAAGCAAGTGTGAGTGG 3540
Qy 3541 CTTACATGGCGATAGCTAGACTGGGCGGTTTATATGGACAGCAAGCAAGCAAGTGTGCGCA 3600
Db 3541 CTTACATGGCGATAGCTAGACTGGGCGGTTTATATGGACAGCAAGCAAGCAAGTGTGCGCA 3600
Qy 3601 GCTGGGCGGCTCTCTGTTAAGTGGGAAAGCCCTGCAAAAGTAAACTGATGGCTTTCTTG 3660
Db 3601 GCTGGGCGGCTCTCTGTTAAGTGGGAAAGCCCTGCAAAAGTAAACTGATGGCTTTCTTG 3660
Qy 3661 CCGCAAGAGTCTGATGGCGCAGGGATCAAGATCTGATCAAGAGACAGGATGAGGATCG 3720
Db 3661 CCGCAAGAGTCTGATGGCGCAGGGATCAAGATCTGATCAAGAGACAGGATGAGGATCG 3720
Qy 3721 TTTGCGCATGATGAAACAGATGGAATTCAGCGAGGCTTCCGGCGCTTGGGTGGAGAGG 3780
Db 3721 TTTGCGCATGATGAAACAGATGGAATTCAGCGAGGCTTCCGGCGCTTGGGTGGAGAGG 3780
Qy 3781 CTATTCGGCTATGATGGGCAACACAGCAATCGGCTCTGATGCGCGCTGTTCCGG 3840
Db 3781 CTATTCGGCTATGATGGGCAACACAGCAATCGGCTCTGATGCGCGCTGTTCCGG 3840
Qy 3841 CTGTACGCGAGGGCGCCCGGCTTCTTTTGTCAAGACCGACCTGTCCGGTCCCTGAAT 3900
Db 3841 CTGTACGCGAGGGCGCCCGGCTTCTTTTGTCAAGACCGACCTGTCCGGTCCCTGAAT 3900
Qy 3901 GAATGCGAGCAGGAGCAGCGGCTATCGTGGCTGGCCACGACGGGCTTCTTGGCGCA 3960
Db 3901 GAATGCGAGCAGGAGCAGCGGCTATCGTGGCTGGCCACGACGGGCTTCTTGGCGCA 3960
Qy 3961 GCTGTGCTCGACGTTGCTCACTGGAAGCGGAAAGGAGCTGGGCTGCTATTTGGGCGAAGTCCG 4020
Db 3961 GCTGTGCTCGACGTTGCTCACTGGAAGCGGAAAGGAGCTGGGCTGCTATTTGGGCGAAGTCCG 4020
Qy 4021 GGGCAGGATCTCTGTCTCATCTCACCTTGTCTCTCGCGAGAAAGTATCCATCATGGCTGAT 4080
Db 4021 GGGCAGGATCTCTGTCTCATCTCACCTTGTCTCTCGCGAGAAAGTATCCATCATGGCTGAT 4080
Qy 4081 GCAATGCGGCGGTGACATCGCTTGTATCCGGCTACCTGCGCATTCGACCAAGCGGAAA 4140
Db 4081 GCAATGCGGCGGTGACATCGCTTGTATCCGGCTACCTGCGCATTCGACCAAGCGGAAA 4140
Qy 4141 CATCGCATCGAGCAGCAGTACTCGGATGGAGCCGCTCTTGTGATCAGGATGATCTG 4200
Db 4141 CATCGCATCGAGCAGCAGTACTCGGATGGAGCCGCTCTTGTGATCAGGATGATCTG 4200
Qy 4201 GACGAAGAGATCAGGGCTCGGCGCAGCGCAACTGTTTGGCAGGCTCAAGCGCGGATG 4260
Db 4201 GACGAAGAGATCAGGGCTCGGCGCAGCGCAACTGTTTGGCAGGCTCAAGCGCGGATG 4260
Qy 4261 CCGACGCGGAGGATCTGCTGTGACCCATGGCGATGCTGCTTGGCGGAATATCATGGTG 4320
Db 4261 CCGACGCGGAGGATCTGCTGTGACCCATGGCGATGCTGCTTGGCGGAATATCATGGTG 4320
Qy 4321 GAAAATGGCGCTTTTCTGGAATTCATGATCTGCGCGGCTGGGTGGTGGCGGACCGCTAT 4380
Db 4321 GAAAATGGCGCTTTTCTGGAATTCATGATCTGCGCGGCTGGGTGGTGGCGGACCGCTAT 4380

| | | | |
|----|------|--|------|
| Qy | 4381 | CAGGACATAGCGTTGGCTACCGTGATATTGCTGAAGAGCTTGGCGGCAATGGCTGAC | 4440 |
| Db | 4381 | | |
| Qy | 4441 | CGCTTCCTCGCTTTACGGTATCGCCGCTCCGATTCGCGAGCGATCGCTTCTATCGC | 4500 |
| Db | 4441 | | |
| Qy | 4501 | CTTCTTGACGAGTTCTTCTGAGCGGGACTCTGGGGTTCGATAAATAAAGATTTATTT | 4560 |
| Db | 4501 | | |
| Qy | 4561 | AGTCTCCAGAAAAAGGGGGGAATGAAGACCCCACTGTAGCTTTGGCAAGCTAGCTTAA | 4620 |
| Db | 4561 | | |
| Qy | 4621 | GTAACGCCATTTTGAAGGCATGGAATAATACATACTGAGAAATAGAGAGTTTCAATCA | 4680 |
| Db | 4621 | | |
| Qy | 4681 | AGGTGAGAAACAGATGGAACAGCTGATATATGGGCCAACAAGGATATCTGTGTAAGCAGT | 4740 |
| Db | 4681 | | |
| Qy | 4741 | TCCTGCCCGGCTCAGGGCCAGAACAGATGGAACAGCTGATATATGGGCCAACAAGGATAT | 4800 |
| Db | 4741 | | |
| Qy | 4801 | TCTGTGTGTAAGCAGTTCTGCCCCGGCTCAGGGCCAGAACAGATGTTCCCAAGATCGG | 4860 |
| Db | 4801 | | |
| Qy | 4861 | TCCAGCCTCAGCAGTTTCTAGAGAACCATCAGATGTTTCCAGGGTCCCCCAAGGACCTG | 4920 |
| Db | 4861 | | |
| Qy | 4921 | AAATGACCTCTGCTTATTTGAACCTAACCAATCAGTTTCCGCTTCTCGCTTCTGTTCCGC | 4980 |
| Db | 4921 | | |
| Qy | 4981 | GCTTCTGCTCCCGAGCTCAATAAAGAGCCCAACCCCTCACTCGGGGGCCAGTCTT | 5040 |
| Db | 4981 | | |
| Qy | 5041 | CCGATTGACTGAGTGGCCGGGTACCGGTGATCCCAATTAACCTCTTGCAGTTGCAATCC | 5100 |
| Db | 5041 | | |
| Qy | 5101 | GACTTGTGGTCTCGCTGTTCTTGGAGGGTCTCTCTGAGTGATTAAGTACCCGTCAGC | 5160 |
| Db | 5101 | | |
| Qy | 5161 | GGGGGTCTTTCATTGG | 5177 |
| Db | 5161 | | |

RESULT 2
AX018983
LOCUS AX018983 5177 bp DNA linear PAT 07-SEP-2000
DEFINITION Sequence 1 from Patent WO9942577.
ACCESSION AX018983
VERSION AX018983.1 GI:10043078
KEYWORDS
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Yla-Herttuala, S., Airenne, K., Kulomaa, M., Marjomaki, V. and Lehtolainen, P.
TITLE Biotin-binding receptor molecules
JOURNAL Patent: WO 9942577-A 1 26-AUG-1999;

| | | |
|---------------------|--|---|
| FEATURES | source | YLA HERTTUALA SEPPO (FI); AIRENNE KARI (FI); KULOMAA MARKKU (FI); MARJOMAKI VARP (FI); LEHTOLAINEN PAULIINA (FI); EUROGENE LIMITED (GB) |
| Location/Qualifiers | 1..5177 | |
| | /organism="synthetic construct" | |
| | /mol_type="unassigned DNA" | |
| | /db_xref="taxon:32630" | |
| | /note="Designed DNA sequence to encode transmembrane protein." | |
| CDS | 1071..2273 | |
| | /note="unnamed protein product" | |
| | /codon_start=1 | |
| | /transl_table=11 | |
| | /protein_id="CAC07722.1" | |
| | /db_xref="GI:10043079" | |
| | /translation="MAQDDFPDQEDTDSCTESVKFDRASVTALLPFPKNGPTLQERMSKYTALITLYLIVPVVLPIYIGVAAQLKWTNCTVGSVNADISPEKGGNGSEDMFPEAVMERMSMESRIQYLSDNENLLDAKNFQNFISITDFNDVLPQLNSLSSIOBHENIIGDISLVGLNTVLDLQFSIETLNGRVQENAFKQOEBMRKLEERIYNASAEIKSIDKQVYLEQRIKGMKLNNITNDLRKLDWEHSQTLKNTILLOQARKSLTGKNTNDLGSNNMTIGAVNSRGEFTGTITATVATSNSEIKESPLHGTQNTINKETQPTFGFTVNWKFSESTTFTGQCFIDRNGKEVLKTMWLURSSVNDIGDDWKATRVGINIFTRLTQKE" | |
| ORIGIN | Query Match 100.0%; Score 5177; DB 6; Length 5177; Best Local Similarity 100.0%; Pred. No. 0; Matches 5177; Conservative 0; Mismatches 0; Indels 0; Gaps 0; | |
| Qy | 1 | TTTGAAGACCCACCCGCTAGGTGGCAAGCTAGCTTAAGTAACGCCCACTTTTGCAGGCAT 60 |
| Db | 1 | TTTGAAGACCCACCCGCTAGGTGGCAAGCTAGCTTAAGTAACGCCCACTTTTGCAGGCAT 60 |
| Qy | 61 | GGAAAAATACATTAAGTGAAGTAAAGTTCAGATCAAGGTCAAGAACAAAGAAACAGC 120 |
| Db | 61 | GGAAAAATACATTAAGTGAAGTAAAGTTCAGATCAAGGTCAAGAACAAAGAAACAGC 120 |
| Qy | 121 | TGAATACCAACAGGATATCTGTGTAAGCGGTTCCTGCCCGGCTCAGGGCCCAAGAAC 180 |
| Db | 121 | TGAATACCAACAGGATATCTGTGTAAGCGGTTCCTGCCCGGCTCAGGGCCCAAGAAC 180 |
| Qy | 181 | GATGAGACAGCTGAGTGATGGGCCAAACAGGATATCTGTGTAAGCAGTTCCTGCCCGG 240 |
| Db | 181 | GATGAGACAGCTGAGTGATGGGCCAAACAGGATATCTGTGTAAGCAGTTCCTGCCCGG 240 |
| Qy | 241 | CTCGGGGCCAAGAACAGATGTTCCCGAGTCCGGTCCAGCCCTCAGCAGTTCCTAGTGAA 300 |
| Db | 241 | CTCGGGGCCAAGAACAGATGTTCCCGAGTCCGGTCCAGCCCTCAGCAGTTCCTAGTGAA 300 |
| Qy | 301 | TCATCAGATGTTTCCAGGGTCCCAAGGACCTGAAATGACCTGTACCTTATTTGAAC 360 |
| Db | 301 | TCATCAGATGTTTCCAGGGTCCCAAGGACCTGAAATGACCTGTACCTTATTTGAAC 360 |
| Qy | 361 | TAAACCAATCAGTTCGCTTCTCGCTTCTGTGTCGGCGCTTCGGCTCCCGAGCTCAATAA 420 |
| Db | 361 | TAAACCAATCAGTTCGCTTCTCGCTTCTGTGTCGGCGCTTCGGCTCCCGAGCTCAATAA 420 |
| Qy | 421 | AGAGCCCAACCCCTCACTCGCGCGGCCAGTCTTCGATAGACTGCGTCCCGGGGTAC 480 |
| Db | 421 | AGAGCCCAACCCCTCACTCGCGCGGCCAGTCTTCGATAGACTGCGTCCCGGGGTAC 480 |
| Qy | 481 | CCGTATTTCCCAATAAGCCTCTTGTGTTTGCATCCGAATCGTGTCTCCGTGTTCTTGG 540 |
| Db | 481 | CCGTATTTCCCAATAAGCCTCTTGTGTTTGCATCCGAATCGTGTCTCCGTGTTCTTGG 540 |
| Qy | 541 | GGAGGGTCTCTCTGAGTGATTGACTACCCACGACGGGGGTCTTTTCAATTTGGGGGCTCGT 600 |
| Db | 541 | GGAGGGTCTCTCTGAGTGATTGACTACCCACGACGGGGGTCTTTTCAATTTGGGGGCTCGT 600 |
| Qy | 601 | CCGGGATTTGGAGACCCCTGCCAGGGACCAACGACCCACGACGGGAGGTAACTGGCC 660 |
| Db | 601 | CCGGGATTTGGAGACCCCTGCCAGGGACCAACGACCCACGACGGGAGGTAACTGGCC 660 |

| | | | |
|----|------|---|------|
| Qy | 661 | AGCAACTTATCTGTCTGTCCGATTGTCTAGTGTCTATGTCTTCAATGTTATGCGCCTGGC | 720 |
| Db | 661 | AGCAACTTATCTGTGTCTGTCCGATTGTCTAGTGTCTATGTCTTGAATGTTATGCGCCTGGC | 720 |
| Qy | 721 | TCTGTACTAGTTAGCTAACTAGCTCTGTATCTGCGCGACCCGTTGGTTCGAACTGACGAGTT | 780 |
| Db | 721 | TCTGTACTAGTTAGCTAACTAGCTCTGTATCTGCGCGACCCGTTGGTTCGAACTGACGAGTT | 780 |
| Qy | 781 | CTGAACACCCGGCCGCAACCCCTGGGAGACGTCCAGGGACTTTTGGGGGCCGTTTTTGTGG | 840 |
| Db | 781 | CTGAACACCCGGCCGCAACCCCTGGGAGACGTCCAGGGACTTTTGGGGGCCGTTTTTGTGG | 840 |
| Qy | 841 | CCGACCTGAGGAGGAGTCGATGTGGAAATCCGACCCCGTCAGGATATGTGGTTCTGGT | 900 |
| Db | 841 | CCCGACCTGAGGAAAGGAGTCGATGTGGAAATCCGACCCCGTCAGGATATGTGGTTCTGGT | 900 |
| Qy | 901 | AGGAGACGAGAACCTTAAACAGTTTCCCGCCTCCGCTCTGGAATTTTGTCTTTCCGTTTGGAA | 960 |
| Db | 901 | AGGAGACGAGAACCTTAAACAGTTTCCCGCCTCCGCTCTGGAATTTTGTCTTTCCGTTTGGAA | 960 |
| Qy | 961 | CCGAAGCGCGCGTCTTGTCTGTGCGACGCAAGCTTGGCGTCGACGTGCGACTCTAGAGGA | 1020 |
| Db | 961 | CCGAAGCGCGCGTCTTGTCTGTGCGACGCAAGCTTGGCGTCGACGTGCGACTCTAGAGGA | 1020 |
| Qy | 1021 | TCAATTCGGCACGAGTAAATCCGGTCTGCCGTCTTTAGGCATATGAAGTATGCGCACGT | 1080 |
| Db | 1021 | TCAATTCGGCACGAGTAAATCCGGTCTGCCGTCTTTAGGCATATGAAGTATGCGCACGT | 1080 |
| Qy | 1081 | GGGATGACTTTTCTGTATCAGCAAGAGGACACTGACAGCTGTACAGAGTCTGTGAAAGTTTCG | 1140 |
| Db | 1081 | GGGATGACTTTTCTGTATCAGCAAGAGGACACTGACAGCTGTGTGAAAGTTTCG | 1140 |
| Qy | 1141 | ATGCTCGCTCAGTGACAGCTTTGCTTCTCCCAATCTTAAATATGCGCCCAACTCTTCAAG | 1200 |
| Db | 1141 | ATGCTCGCTCAGTGACAGCTTTGCTTCTCCCAATCTTAAATATGCGCCCAACTCTTCAAG | 1200 |
| Qy | 1201 | AGAGGATGAAGTCTTTATAAACCTGCACGTATCACCCCTTATCTCATTTGCTTTGTAGTTC | 1260 |
| Db | 1201 | AGAGGATGAAGTCTTTATAAACCTGCACGTATCACCCCTTATCTCATTTGCTTTGTAGTTC | 1260 |
| Qy | 1261 | TCGTGCCCATCATTTGGCATATAGTGGCAGCTCAGCTCTCGAAATGGGAAACGAAAGAAATTCGA | 1320 |
| Db | 1261 | TCGTGCCCATCATTTGGCATATAGTGGCAGCTCAGCTCTCGAAATGGGAAACGAAAGAAATTCGA | 1320 |
| Qy | 1321 | CGGTTGGCTCAGTTAATGCGAGATATATCTCCAAGTCGGAAGGCAAGGAAATGGCAGTG | 1380 |
| Db | 1321 | CGGTTGGCTCAGTTAATGCGAGATATATCTCCAAGTCGGAAGGCAAGGAAATGGCAGTG | 1380 |
| Qy | 1381 | AAGATGAATGAGATTTTCGAGAGAGCTCTGATGGAACCATGAGCAACATGGGAAGCAGAA | 1440 |
| Db | 1381 | AAGATGAATGAGATTTTCGAGAGAGCTCTGATGGAACCATGAGCAACATGGGAAGCAGAA | 1440 |
| Qy | 1441 | TCCAGTATCTTTTCAGATAATGAAGCCCAATCTCCTAGATGCTTAAGAAATTTCCAAAAATTTCA | 1500 |
| Db | 1441 | TCCAGTATCTTTTCAGATAATGAAGCCCAATCTCCTAGATGCTTAAGAAATTTCCAAAAATTTCA | 1500 |
| Qy | 1501 | GCATACAACTGATCAAAAGATTTAATGATGTTCTTTTCCAGTCAAAATCTCTTACCTTCCT | 1560 |
| Db | 1501 | GCATACAACTGATCAAAAGATTTAATGATGTTCTTTTCCAGTCAAAATCTCTTACCTTCCT | 1560 |
| Qy | 1561 | CCATCCAGGAACATGAGAAATATCATAGGGGATATCTCCAAGTCAATAGTAGGTCTGAACA | 1620 |
| Db | 1561 | CCATCCAGGAACATGAGAAATATCATAGGGGATATCTCCAAGTCAATAGTAGGTCTGAACA | 1620 |
| Qy | 1621 | CCACAGTACTTGATTTTCAGTTCAGTATTTGAAACACTGAAATGGCAGAGTCCACAGAGAAATG | 1680 |
| Db | 1621 | CCACAGTACTTGATTTTCAGTTCAGTATTTGAAACACTGAAATGGCAGAGTCCACAGAGAAATG | 1680 |
| Qy | 1681 | CATTTAACAACAAGAGGAGATGCGTAAATTTAGAGGAGCGTATATACATGCAATCAGCAG | 1740 |
| Db | 1681 | CATTTAACAACAAGAGGAGATGCGTAAATTTAGAGGAGCGTATATACATGCAATCAGCAG | 1740 |

| | | | |
|----|------|--|------|
| QY | 1741 | AAATTTAAGTCTCTAGATGNAACAAACAAAGTATATTTGGAAACGGAAATAAAGGGGGAATGA | 1800 |
| DB | 1741 | AAATTTAAGTCTCTAGATGNAACAAACAAAGTATATTTGGAAACGGAAATAAAGGGGGAATGA | 1800 |
| QY | 1801 | AACTGTGCAATAATACATAATATGATCTCAGGCGTGAAGGATTCGGGAACATTTCTCAGACAT | 1860 |
| DB | 1801 | AACTGTGCAATAATATCACTAATATGATCTGAGGCTGAAGGATTCGGGAACATTTCTCAGACAT | 1860 |
| QY | 1861 | TGAAAAATATCACTTTTACTCCAAAGTGCAGAAAAGTCTCGTGACTGGGAAATGGACCA | 1920 |
| DB | 1861 | TGAAAAATATCACTTTTACTCCAAAGTGCAGAAAAGTCTCGTGACTGGGAAATGGACCA | 1920 |
| QY | 1921 | ACGATCTGGGCTCCAAACATGACCATCGGGCTGTGAAACAGCAGAGTGTAATTCACAGGCA | 1980 |
| DB | 1921 | ACGATCTGGGCTCCAAACATGACCATCGGGCTGTGAAACAGCAGAGTGTAATTCACAGGCA | 1980 |
| QY | 1981 | CCTACATCACAGCCGTAAACAGCCACATCAAAATGAGATCAAAGAGTCACCACTGCATCGGA | 2040 |
| DB | 1981 | CCTACATCACAGCCGTAAACAGCCACATCAAAATGAGATCAAAGAGTCACCACTGCATCGGA | 2040 |
| QY | 2041 | CACAAAAACCATCAACAGAGGACCCAGCCACCTTTGGCTTCCCGTCAATTTGGAAGT | 2100 |
| DB | 2041 | CACAAAAACCATCAACAGAGGACCCAGCCACCTTTGGCTTCCCGTCAATTTGGAAGT | 2100 |
| QY | 2101 | TTTCAGAGTCACCACTGTCTTCA CGGGCCAGTCTTCA TAGACAGGAATGGGAGGAGG | 2160 |
| DB | 2101 | TTTCAGAGTCACCACTGTCTTCA CGGGCCAGTCTTCA TAGACAGGAATGGGAGGAGG | 2160 |
| QY | 2161 | TCCTGAAGCACAATGGCTGCTCGGCTCAAGTGTAAATGACATTTGGTGATGACTGGAAAG | 2220 |
| DB | 2161 | TCCTGAAGCACAATGGCTGCTCGGCTCAAGTGTAAATGACATTTGGTGATGACTGGAAAG | 2220 |
| QY | 2221 | CTACCAGGCTCGGCATCAACATCTTCACTCGCTCGCCACACAGAGGAGTGAAGTGAAG | 2280 |
| DB | 2221 | CTACCAGGCTCGGCATCAACATCTTCACTCGCTCGCCACACAGAGGAGTGAAGTGAAG | 2280 |
| QY | 2281 | ACCAAGGCTCTCTCGGACTCAAGTGTAAAGGAGATAGAGGCCCTCTCTGGAACAAATGG | 2340 |
| DB | 2281 | ACCAAGGCTCTCTCGGACTCAAGTGTAAAGGAGATAGAGGCCCTCTCTGGAACAAATGG | 2340 |
| QY | 2341 | TATACAGGCTTTCCAGGCTCAATAGGTACTCCAGTCTTAAAGGTGATCGGGGGATCT | 2400 |
| DB | 2341 | TATACAGGCTTTCCAGGCTCAATAGGTACTCCAGTCTTAAAGGTGATCGGGGGATCT | 2400 |
| QY | 2401 | CTGGTTTACCTGGAGTTCCAGGATTTCCAGGACCAATGGGGAAGACCGGGAAGCCAGGAC | 2460 |
| DB | 2401 | CTGGTTTACCTGGAGTTCCAGGATTTCCAGGACCAATGGGGAAGACCGGGAAGCCAGGAC | 2460 |
| QY | 2461 | TTAATGGACAAAAAGCCAGGAAGGGAGAAAAAGGGAGTGGAAAGCAATCTTA | 2520 |
| DB | 2461 | TTAATGGACAAAAAGCCAGGAAGGGAGAAAAAGGGAGTGGAAAGCAATCTTA | 2520 |
| QY | 2521 | ATACAGTCCGACTGGTGGTGGCAGGGCCCTCACGAAGCAGATGGAGATTTTTCACG | 2580 |
| DB | 2521 | ATACAGTCCGACTGGTGGTGGCAGGGCCCTCACGAAGCAGATGGAGATTTTTCACG | 2580 |
| QY | 2581 | AAGGCCAGTGGGGTACGGTGTGTGACGACCGCTGGGAACTGCGTGAGGAGCTGGTCTCT | 2640 |
| DB | 2581 | AAGGCCAGTGGGGTACGGTGTGTGACGACCGCTGGGAACTGCGTGAGGAGCTGGTCTCT | 2640 |
| QY | 2641 | GCAGGAGCTTGGGATACAAAGGTGTTCAAAGTGTGCAATAAGCAGCTTATTTTGGAAAAAG | 2700 |
| DB | 2641 | GCAGGAGCTTGGGATACAAAGGTGTTCAAAGTGTGCAATAAGCAGCTTATTTTGGAAAAAG | 2700 |
| QY | 2701 | GTACGGGTCCAATATGGCTGAATGAAGTATTTTGTTCGGGAAGAGTCAATCCATTGAAG | 2760 |
| DB | 2701 | GTACGGGTCCAATATGGCTGAATGAAGTATTTTGTTCGGGAAGAGTCAATCCATTGAAG | 2760 |
| QY | 2761 | AGTGCAAGAAATAGACAGTGGGGTGTGAGAGCCTGTTCCGACGACGAAGATGCTGGGGGTC | 2820 |
| DB | 2761 | AGTGCAAGAAATAGACAGTGGGGTGTGAGAGCCTGTTCCGACGACGAAGATGCTGGGGGTC | 2820 |
| QY | 2821 | ACTTTGCACCTACATAATATGCATCATATTTTCAATTCATTTTTTAACTGTGTTAAAGTG | 2880 |

| | | | |
|----------------------------|------|---|------|
| Qy | 5041 | CCGATTGACTGAGTCGCGCCGGGTACCCGCTGTATCCAAATAAACCCCTCTTGAGTTGCATCC | 5100 |
| Db | 5041 | CCGATTGACTGAGTCGCGCCGGGTACCCGCTGTATCCAAATAAACCCCTCTTGAGTTGCATCC | 5100 |
| Qy | 5101 | GACTTGTGGTCTCGCTGTTCTCTTGGAGGGGTCTCCTCTGAGTGATTGACTTACCCGTCAGC | 5160 |
| Db | 5101 | GACTTGTGGTCTCGCTGTTCTCTTGGAGGGGTCTCCTCTGAGTGATTGACTTACCCGTCAGC | 5160 |
| Qy | 5161 | GGGGGCTTTTCATTTGG | 5177 |
| Db | 5161 | GGGGGCTTTTCATTTGG | 5177 |
| RESULT 3 | | | |
| A80474 | | | |
| LOCUS | | | |
| DEFINITION | | | |
| ACCESSION | | | |
| VERSION | | | |
| KEYWORDS | | | |
| SOURCE | | | |
| ORGANISM | | | |
| REFERENCE | | | |
| AUTHORS | | | |
| TITLE | | | |
| JOURNAL | | | |
| FEATURES | | | |
| source | | | |
| ORIGIN | | | |
| Query Match | | | |
| Best Local Similarity | | | |
| Matches 3339; Conservative | | | |
| Qy | | | |
| Db | | | |
| Qy | | | |
| Db | | | |
| Qy | | | |
| Db | | | |
| Qy | | | |
| Db | | | |
| Qy | | | |
| Db | | | |
| Qy | | | |
| Db | | | |
| Qy | | | |
| Db | | | |
| Qy | | | |
| Db | | | |
| Qy | | | |
| Db | | | |
| Qy | | | |
| Db | | | |
| Qy | | | |
| Db | | | |
| Qy | | | |
| Db | | | |
| Qy | | | |
| Db | | | |
| Qy | | | |
| Db | | | |
| Qy | | | |
| Db | | | |
| Qy | | | |
| Db | | | |
| Qy | | | |
| Db | | | |
| Qy | | | |
| Db | | | |
| Qy | | | |
| Db | | | |
| Qy | | | |
| Db | | | |
| Qy | | | |
| Db | | | |
| Qy | | | |
| Db | | | |
| Qy | | | |
| Db | | | |
| Qy | | | |
| Db | | | |
| Qy | | | |
| Db | | | |
| Qy | | | |
| Db | | | |
| Qy | | | |
| Db | | | |
| Qy | | | |
| Db | | | |
| Qy | | | |
| Db | | | |
| Qy | | | |
| Db | | | |
| Qy | | | |

Db 1727 GCTGGGAGACATCTATCGTGCATCTCTTTGAATGGCGCTGGCGGACATTCGACTGGAG 1786
Qy 1619 CACCACAGTACTGATTTGTCAGTTTCAGTATTGAAACACTGAATGGCAGAGTCCCAAGAGAA 1678
Db 1787 TGGGAACACTAATTAGCTCTTAATGGGTTTGGAGAGTTTCAGGTTTCTCTCCAAATGAA 1846
Qy 1679 TGCATTTAAACAAACAGAGAGAGATCGCTAAATTTAGAGGAGGCTATATACAAATGCATCAGC 1738
Db 1847 AACATTGTCTACTTAATCCGAAACAGGCCCTGGTGGAAAGATACCAGCCCATCAGCTAC 1906
Qy 1739 AGAAATTAAGTCTCTAGATGAAACAAAGTATATTTGGAAACAGGAATAAAGGGGAAAT 1798
Db 1907 AAGATCTGCAGTCGATCGGCAATCGGCAATGAAATGAAATTCAGAGACATGGTGACCAATGCAAC 1966
Qy 1799 GAAACTGTTGAATATACATATGATCTGAGGCTGAGGATTTGGGAACATTTCTCAGAC 1858
Db 1967 AATGTTGGAG--TTGCTATTTATGTGGATGCTGTGTCAATCACATGTGTGGATCTATGGG 2025
Qy 1859 ATTGAAATAATACATTTACTTCAAGGTGCCAGAAAGTGTCTCGCTGACTGGGAATGGAC 1918
Db 2026 TGGCAGGGCACCCACTCAACATGTGGAGCTATTTCAACACCGGACTAGAGATTTTCC 2085
Qy 1919 CAACGATCTGGGCTCCAAACATGACATCGGGGCTGTGAACAGCAGAGGTGAATTCACAGG 1978
Db 2086 CGCTGTGCGTACTCTGCTCGGATTTTCAATGACGGCAAAATGTCACTGCAAGTGGAGA 2145
Qy 1979 CACCTACATCACGCGTAAACAGCCACATCAATAGAGATCAAGAGTCAACACATGCATGG 2038
Db 2146 CATCGAAATTTATGGGGACATGTATCAGGTCCGGGATTTGCAAGTTGTCCAGCCTTCTTGA 2205
Qy 2039 GACA--CAAAAACCAATCAACAGAGGACCCAGCCACCTTTGGCTTCACCGTCAATTTGA 2097
Db 2206 TCTGGCTCTGGAGAGGACTATGTACGCTCAACAAATTCGACGCTACATGAATCACCTCAT 2265
Qy 2098 AGTTTTCAGAGTCCACCATCTCTTTCACGGGCCAGTGTCTTCAATAGACAGGAATGGGAAGG 2157
Db 2266 TGATATGGGTAGCAGGGTTCCGGATCGATGCTGCCAAGCATATGTGGCCAGGGGACAT 2325
Qy 2158 AGTCTCTGAAGACCATGTGGCTGTGGTCAAGTGTAAATGACATTTGGTGTGATGACTGGA 2217
Db 2326 AAGAGCGTTTCTGGACAAACTGACAGATCTAAATACATCAGTGGTTTTTCAGCAGGAACGAA 2385
Qy 2218 AAGTCTACAGGGTC--GGCATCAACATCTTCACTCGCCTGCGCACACAGAGGAGTGAGTG 2276
Db 2386 ACCCTTTATTACCAAGAGTAAATTGACTTTGGAGGAGAGCCATCAACAGCACTCAGTA 2445
Qy 2277 AGTCAACAGGTCTCTCTG---GACTCCAGGTGAAAGGAGATAGAGGCCCTCTCGAC 2333
Db 2446 CTTTGGGAATGGCGCGGTGACAGAAATCAAGTATGTGTGTCCAAACTGGGGACGGTATCCG 2505
Qy 2334 AAAATGGTATACAGGCTTTCAGGCTCTAATAGGTACTCCAGGTCTTAAAGGTGATCGGG 2393
Db 2506 GAAGTGAATGGAGAGATAGGCTACTTAAAGAACTGGGGAGAAAGGCTGGGGCTTTGT 2565
Qy 2394 GGGATCTCTGTTTACCTGGAGTTTCGAGATTTCCAGGAC--CAATGGGGAAGACCCGGA 2451
Db 2566 GCCTTCTGACAGAGCCCTCTGTTTGTGGATTAACACAGCAACCAACGCGGGGGCAGGGGC 2625
Qy 2452 AGCCAGGACTTAATGGAACAAAGGCCAGAGGGAGAAAGAGGAGTGAAGCATGCAAA 2511
Db 2626 AGCGGAGCTTCCATCTTTACTTTCTGGATGCGCAGGCTTTTAAATAGGCGGTGGTTTT 2685
Qy 2512 GACAACTAATACAGTCCGACTGGTGGTGGTGGCAGGGCCCTCACGAAGGCAGAGTGAGA 2571
Db 2686 CATGCTCGCTCATCCGTAACGGGTTTCAACGGGTGATGTCAAGTTTATCGTTGGCCAAAGATA 2745
Qy 2572 TTTTTCAGGAAGGCCAGTGGGTTACGGTGTGTGACGACCGCTGGGAACTGGGTGGAGGAC 2631
Db 2746 TTTTGGAAACCGAGTGGATGTATACGACTGGGTGGGACCAACCAAGTAACTCGGACGATC 2805
Qy 2632 T----GGTCTGTCCAGGAGCTTGGGATACAAAGGTGTTCAAAGGTGTTCAGGTGCAATAACGAGCT 2687
Db 2806 GACGAAGTCCGTTTCAATCAACGACAGACACTACCTGTGGCAATGACTGGGTCTCGGAACA 2865

Qy 2688 TATTTTGGAAAAAGTACGGGTCCAATATGGCTGAATGAAGTATTTTGTTCGGGAAGAG 2747
Db 2866 TCGTGGCGCAAAATAGG----AACATGGTTATCTTCCGTAATGTGGTAGACGGTCAGC 2921
Qy 2748 TCATCCATTTGAAGAGTGCAGAAATTAGACAGTGGGGTGTGAGAGCCTGTTCSCACGACGAA 2807
Db 2922 CTTTCTCAAACTGGTGGGACAAACGGGAGCAATCAAGTAGCTTTTGGTGGCGGCGACAGAG 2981
Qy 2808 GATGCTGGGGTCACTTTTGGCACCTATACATATGCAATATTTTTCATTCACATATTTTAA 2867
Db 2982 GCTTCATTGTCTTTAAATGAATGATGACTGGTATATGAATTCGATTTTGCAACTGGTCTGC 3041
Qy 2868 CTGTTATAAAGTGAATTTTTCCTTTGCTTTCACATAAATCAGCTTAATTAATAATTAAGA 2927
Db 3042 CTGCTGAAACCTACTCGGATGTTTATTTCTGGACAAAGGAAGGAGCGGTGCTGACTGAA 3101
Qy 2928 AACTAAGAAATTTTATCCACAGAAAGGAATATTTTAAATAATCCTGGAATAACATATAAA 2987
Db 3102 AGCAGGTGTACGTTTCTCGATGGAAGGCCAATTTCCAGATTTAGTAAACAGCGATGAAG 3161
Qy 2988 TAGCTTCATATTTGCTTCAAAATACCAGAACATTTCACTTCTCTAGGTTTAAAGTGGC 3047
Db 3162 ATCCATTTGTGCAATTTACGTTTGAATGSCCAAGTTATAAGCTTCGAGGATCCACTAGTAAC 3221
Qy 3048 TCGTGGCGAAATTCATCCCTCAGGATATAGTAGTTTCGCTTTTGTGCAATAGGAGGGGAAA 3107
Db 3222 GGGCGCAGTGTCTGGAATTCG-----CTTGTGACATCTAGGGCGGCNAATTCGCC 3276
Qy 3108 TGTAGTCTTATGCAATACTCTTGTAGTCTTTCACACATGGTAAACGATGAGTTAGCAACATG 3167
Db 3277 CCTCTCCCCCCCCCTTAACGTTTACTTGGCGGAAGCGCTTGGAAATAAGGCCGTGTGTG 3336
Qy 3168 CTTTAAAGAGAGAAAGAACGACCGTGCATGCCGATTTGGTGGGAAGTAAAGTGGTACGATC 3227
Db 3337 TTTGTCTATATGTGATTTTCCACCATATTCGCGTCTTTTGGCAA-----TGTGAGGGCC 3390
Qy 3228 GTGCTTTATTAGGAAGCAACAGAGCGGTCTGACATGGAATGGAGCAACCACTCAATTC 3287
Db 3391 CGAAACCTGGCCCTGTCTTTTGACGAGATTTCTTAGGGGTCTTTCCCTCTCGCCAA 3450
Qy 3288 GCATTCAGAGATATTTGATTTTAAAGTGCCTAGTCTGATCAGCAAAACGCCATTTGACCAT 3347
Db 3451 GGAATGCAAGTCTGTTGAATGCTGTAAGAGAGCAGTCTCTCTGGAAGCTTCTTGAAGA 3510
Qy 3348 TCACCAATTTGGTGTGCACTCCAGCTTCAAGTTCACGCTGCCGACAGCAGTACAGGGCCAGGG 3407
Db 3511 CAACCAACGCTCTGTAGCGACCTTTTGACGACGCGGAACCCCCACCTGGCGACAGGTGC 3570
Qy 3408 CTGCTAAAGGAAGCGGAACACGCTAGAAAGCGAGTCCGACAGAAACGGTGTGACCCCGAT 3467
Db 3571 CTCTGGGCCCAAGCCACGCTGATATAGATACACTGAAAGGGCGGCAACCCAGTGC 3630
Qy 3468 GAATGTCACTACTGGGCTATCTGGAACA-AGGGAACACGCAAGCGCAAGAGAAAGCAGG 3526
Db 3631 CACGTTGTGAGTTGGATAGTTGTGGAAGAGTCAATGGCTCTCTCAAGGTAGTCAAC 3690
Qy 3527 TAGCTTGCAGTGGGCTTACATGGCGATAGCTAGACTGGGCGGTTTTTATGACAGCAAGCG 3586
Db 3691 AAGGGGTGAAGGATGCCAGAAAGTACCCCATTTGTATGGGAATCTGATCTGGGGCCCTCG 3750
Qy 3587 AACCGGAATTCACAGCTGGGGCGCCCTCTGTAAGGTTGGGAAGCCCTGCAAAAGTAAACT 3646
Db 3751 GTGCAATGCTTTTACATGTGTTTGTAGTCAGGTTTAAAAAGCTTAGGCCCCCCCCGAACAC 3810
Qy 3647 GGAATGGCTTTTTCGCCCAAGGATCTGATGGCGCAGGGATCAAGATCTGATCAAGAGA 3706
Db 3811 GGGGACGTGTTTTTCTTTTGAAGAAACACGATGATAAGCTTGGCCACAAACCCGGGATAATTC 3870
Qy 3707 CAGATGAGGATCGTTTTCGATGATTTGAACAAGATGGAATTCACGAGAGTTTCTCCGGCG 3766
Db 3871 CTGACGCAATATATGGGATCGGCCAATTTGAACAAGATGGAATTCACGAGGTTTCTCCGGCCG 3930

3767 CTTGGTGGAGAGGCTATTTCGGCTATGACTGGGCAACAACAGCAATCGGCTGCTCTGATG 3826
Db CTTGGTGGAGAGGCTATTTCGGCTATGACTGGGCAACAACAGCAATCGGCTGCTCTGATG 3990
3827 CCGCGGTGTTCCGGCTGTGACGCGAGGGGCGCCGGGTTCTTTTGTCAAGACCGACCTGT 3886
Db CCGCGGTGTTCCGGCTGTGACGCGAGGGGCGCCGGGTTCTTTTGTCAAGACCGACCTGT 4050
3887 CCGGTGCCCTGAATGAATGACGAGAGGAGGCGGGCTATCTGTGGTGGCCACGACGG 3946
Db CCGGTGCCCTGAATGAATGACGAGAGGAGGCGGGCTATCTGTGGTGGCCACGACGG 4110
3947 GCGTTCTTGGCGAGCTGTCTCGACCTTGTCACTGAAGCGGGAAGGACTGCGTCTGCTAT 4006
Db GCGTTCTTGGCGAGCTGTCTCGACCTTGTCACTGAAGCGGGAAGGACTGCGTCTGCTAT 4170
4007 TGGGCGAAGTGC CGGGGCGAGGATCTCTGTCTCATCTCACCTTGTCTTCCGCGAGAAAGTAT 4066
Db TGGGCGAAGTGC CGGGGCGAGGATCTCTGTCTCATCTCACCTTGTCTTCCGCGAGAAAGTAT 4170
4067 CCATCATGGCTGATGCAATGCGGGGCTGCAATACGCTTGATCGGGCTACCTGCCCATTCG 4126
Db CCATCATGGCTGATGCAATGCGGGGCTGCAATACGCTTGATCGGGCTACCTGCCCATTCG 4230
4127 ACCACCAAGCGAATCGCATCGAGCGACACTCTCGATCGAAGCGGCTCTGTCG 4186
Db ACCACCAAGCGAATCGCATCGAGCGACACTCTCGATCGAAGCGGCTCTGTCG 4350
4187 ATCAGGATGATCTCGACGAAGAGCATCAGGGGCTCGCGCGACCGCAACTGTTGCCAGGC 4246
Db ATCAGGATGATCTCGACGAAGAGCATCAGGGGCTCGCGCGACCGCAACTGTTGCCAGGC 4410
4247 TCAAGGCGCGCATGCCCGAGCGGAGGATCTGCTGTGAGCCCATGCGGATGCGCTGCTTC 4306
Db TCAAGGCGCGCATGCCCGAGCGGAGGATCTGCTGTGAGCCCATGCGGATGCGCTGCTTC 4470
4307 CGAATATCATGGTGGAAATGCGCGCTTTCTGGAATTCATCGACTGCGCGGCTGGTG 4366
Db CGAATATCATGGTGGAAATGCGCGCTTTCTGGAATTCATCGACTGCGCGGCTGGTG 4530
4367 TGGCGGACCGCTATCAGACATACGCTTGGCTACCGTGATATGCTGAAGAGCTTGGCG 4426
Db TGGCGGACCGCTATCAGACATACGCTTGGCTACCGTGATATGCTGAAGAGCTTGGCG 4590
4427 GCGAATGGGCTGACCGCTTCTCTGTCTTTACGGTATCGCGCTCCCGATTCGCGAGCGCA 4486
Db GCGAATGGGCTGACCGCTTCTCTGTCTTTACGGTATCGCGCTCCCGATTCGCGAGCGCA 4590
4487 TCGCCTTCTATCGCCTTCTTGAAGAGTCTTCTGAGCGGAGCTCTGGGGTTCGATAAAT 4546
Db TCGCCTTCTATCGCCTTCTTGAAGAGTCTTCTGAGCGGAGCTCTGGGGTTCGATAAAT 4710
4547 AAAAGATTTTATTTAGTCTCCAGAAAGGGGGAATGAAGACCCCACTGTAGTTTG 4606
Db AAAAGATTTTATTTAGTCTCCAGAAAGGGGGAATGAAGACCCCACTGTAGTTTG 4770
4607 GCAAGCTAGCTTAAGTAACGCATTTCGCAAGGCGATGGAATAATACATACTGAGATAG 4666
Db GCAAGCTAGCTTAAGTAACGCATTTCGCAAGGCGATGGAATAATACATACTGAGATAG 4830
4667 AGAATGTCAGATCAAGGTGAGGAACAGATGAGCAAGCTGATATGCGGCCAAACAGGATAT 4726
Db AGAATGTCAGATCAAGGTGAGGAACAGATGAGCAAGCTGATATGCGGCCAAACAGGATAT 4890
4727 CTGTGGTAAGCAGTTCCTGCCCGGCTCAGGGCCAAACAGATGGAAGCTGATATG 4786
Db CTGTGGTAAGCAGTTCCTGCCCGGCTCAGGGCCAAACAGATGGAAGCTGATATG 4950
4787 GGCCAAACAGGATATCTGTGTAGCAGTTCCTGCCCGGCTCAGGGCCAAACAGATG 4846
Db GGCCAAACAGGATATCTGTGTAGCAGTTCCTGCCCGGCTCAGGGCCAAACAGATG 5010
4847 GTCCCCAGATGCGGCTCCAGCCCTCAGCAGTTTCTAGAGAACCATCAGATGTTTCCAGGGT 4906

5011 GTCCCCAGATGCGGTCAGGCCCTCAGCAGTTTCTAGAGAACCATCAGATGTTTCCAGGGT 5070
Qy GTCCCCAGATGCGGTCAGGCCCTCAGCAGTTTCTAGAGAACCATCAGATGTTTCCAGGGT 5070
4907 GCGCCAAAGACCTGAAATGACCTGTGCTTATTTGAACCTAACCAATCAGTTCGCTTCTC 4966
Db GCGCCAAAGACCTGAAATGACCTGTGCTTATTTGAACCTAACCAATCAGTTCGCTTCTC 5130
4967 GCTTCTGTTCCGGGCTTCTGCTCCCGAGGCTCAATAAAGAGCCCAACCCCTCACTC 5026
Db GCTTCTGTTCCGGGCTTCTGCTCCCGAGGCTCAATAAAGAGCCCAACCCCTCACTC 5190
5027 GGGGCGCAGTCTCCGATGACTAGTCCCGGGTACCGTGTATCCATAAACCCCTC 5086
Qy GGGGCGCAGTCTCCGATGACTAGTCCCGGGTACCGTGTATCCATAAACCCCTC 5086
Db GGGGCGCAGTCTCCGATGACTAGTCCCGGGTACCGTGTATCCATAAACCCCTC 5250
5087 TTGCAGTTGTCATCGACTTGTGCTCTGCTTCTTGGGAGGCTCTCTCTGAGTGATT 5146
Db TTGCAGTTGTCATCGACTTGTGCTCTGCTTCTTGGGAGGCTCTCTCTGAGTGATT 5310
5147 GACTACCGCTCAGCGGGGCTTTTCATTGG 5177
Db GACTACCGCTCAGCGGGGCTTTTCATTGG 5341

RESULT 4
BD136838
LOCUS
DEFINITION
BD136838
Transsonomats associated with gene transfer into mammary epithelial cells.
ACCESSION
BD136838.1 GI:23231783
KEYWORDS
JP 2002504365-A/1.
SOURCE
unidentified
ORGANISM
unclassified.
REFERENCE
1 (bases 1 to 7699)
AUTHORS
Gavora,J.S., Falconer,M.M., Nguyen,T.H. and Benkel,B.F.
TITLE
Transsonomats associated with gene transfer into mammary epithelial cells
JOURNAL
Patent: JP 2002504365-A 1 12-FEB-2002;
HER MAJESTY IN RIGHT OF CANADA AS REPRESENTED BY THE MINISTER OF
AGRICULTURE AND AGRI FOOD CANADA
COMMENT
OS Unidentified
PN JP 2002504365-A/1
PD 12-FEB-2002
PF 25-JUN-1998 JP 2000533535
PR 24-FEB-1998 CA 2224108
PI JAN S GAVORA,MARCIA M FALCONER,THUY H NGUYEN,BERNHARD F BENKEL
PC A01K67/027,C12N9/64,C12N15/09,C12P21/00,C12N15/00 CC
Strandedness: Double;
CC Topology: Circular;
Transsonomats associated with gene transfer into mammary CC
epithelial cells
FH Key Location/Qualifiers
FT source 1..7699
FT /organism='Unidentified'.
FEATURES
source
1..7699
/organism='unidentified'
/mol_type='genomic DNA'
/db_xref='taxon:32644'

ORIGIN
Query Match 40.2%; Score 2078.8; DB 6; Length 7699;
Best Local Similarity 64.3%; Pred. No. 0;
Matches 3339; Conservative 0; Mismatches 1814; Indels 38; Gaps 14;
Qy 1 TTTTGAAGAGCCCCCAGCGTAGGTGGCAAGCTAGCTTAAGTAACCGCACTTTTGAAGGCAT 60
Db 175 TTTTGAAGAGCCCCCAGCGTAGGTGGCAAGCTAGCTTAAGTAACCGCACTTTTGAAGGCAT 234
Qy 61 GGAATAATACATTAATCTGAGTAATAGAAAAGTTTCAGATCAAGGTCAGGAACAAAGAACAGC 120

235 GGAATAATACATACTGAGATAGAAAAGTTTCAGATCAAGGTCAGGAACAAAGAAACAGC 294
121 TGAATACCAACAGGATATCTGTGTAGCGGTTCTGCCCGGCTCAGGCCCAAGAAC 180
295 TGAATACCAACAGGATATCTGTGTAGCGGTTCTGCCCGGCTCAGGCCCAAGAAC 354
181 GATGAGACAGCTGAGTGATGGGCCCAACAGGATATCTGTGTAGCGGTTCTGCCCGG 240
355 GATGAGACAGCTGAGTGATGGGCCCAACAGGATATCTGTGTAGCGGTTCTGCCCGG 414
241 CTCGGGGCAAGAACAGATGTTCCCGAGATGCGGTCCAGCCCTCAGAGTTCTTAGTGAA 300
415 CTCGGGGCAAGAACAGATGTTCCCGAGATGCGGTCCAGCCCTCAGAGTTCTTAGTGAA 474
301 TCATCAGATGTTCCAGGGTCCCGCAAGGACCTGAAATGACCCCTGTACCTTATTTGAAC 360
475 TCATCAGATGTTCCAGGGTCCCGCAAGGACCTGAAATGACCCCTGTACCTTATTTGAAC 534
361 TAAACCAATCAGTTCTCGCTTCTCGCTTCTGTTCGCGGCTTCCGCTCTCCGAGCTCAATAA 420
535 TAAACCAATCAGTTCTCGCTTCTGTTCGCGGCTTCCGCTCTCCGAGCTCAATAA 594
421 AGAGCCCAACCCCTCACTCGGCGGCCAGTCTTCGAGTAGACTGCGTCCCGGGTAC 480
595 AGAGCCCAACCCCTCACTCGGCGGCCAGTCTTCGAGTAGACTGCGTCCCGGGTAC 654
481 CCGTATTTCCCAATAAGGCTCTTCTGTTTCGATCCGAACTGCTGCTTCGCTGTTCTG 540
655 CCGTATTTCCCAATAAGGCTCTTCTGTTTCGATCCGAACTGCTGCTTCGCTGTTCTG 714
541 GGAGGGTCTCTCTGAGTGATGTAACCCACGACGGGGTCTTTCATTTGGGGGCTCGT 600
715 GGAGGGTCTCTCTGAGTGATGTAACCCACGACGGGGTCTTTCATTTGGGGGCTCGT 774
601 CCGGGATTTGAGACCCCTGCCCCAGGACACCGAACCCACCGAGGAGTAAGCTGGCC 660
775 CCGGGATTTGAGACCCCTGCCCCAGGACACCGAACCCACCGAGGAGTAAGCTGGCC 834
661 AGCAACTTATCTGTCTGTCGATGCTAGTCTGATCTGATCTGATCTGATCTGATCTG 720
835 AGCAACTTATCTGTCTGTCGATGCTAGTCTGATCTGATCTGATCTGATCTGATCTG 894
721 TCTGTACTAGTTAGTAACTAGCTCTGTATCTGCGGACCCGCTGGTGAACCTGACGAGTT 780
895 TCTGTACTAGTTAGTAACTAGCTCTGTATCTGCGGACCCGCTGGTGAACCTGACGAGTT 954
781 CTGAACACCCCGCGCAACCCCTGCGGAGCTCCAGGACCTTTGCGGCGCTTTTGTGG 840
955 CTGAACACCCCGCGCAACCCCTGCGGAGCTCCAGGACCTTTGCGGCGCTTTTGTGG 1014
841 CCGACCTGAGGAAGGAGTCGATGTGGATCCGACCCCGCTCAGGATATGTTGTTCTGTT 900
1015 CCGACCTGAGGAAGGAGTCGATGTGGATCCGACCCCGCTCAGGATATGTTGTTCTGTT 1074
901 AGGAGACGAGAACCTTAAACAGTTTCCCGCTCTCGTCTGAAATTTTGTCTTGGTTTGGAA 960
1075 AGGAGACGAGAACCTTAAACAGTTTCCCGCTCTCGTCTGAAATTTTGTCTTGGTTTGGAA 1134
961 CCGAAGCGCGCGCTTGTCTGCTGACAGCCCAAGCTTGGGCTGCAAGTCTCTAGAGGA 1020
1135 CCGAAGCGCGCGCTTGTCTGCTGACAGCGCTGAGCATCTGCTGTGTTGTTCTCTGCT 1194
1021 TCAATTCCGACGAGTAATTCGGTGTGCTGCTCTTTAGGACATATGAAATGACGACAGT 1080
1195 GACTGTGTTCTGATTTGTCTGAAAAATTAGGGCCAGACTGTTTACCACTCCCTTAAGTTT 1254
1081 GGGATGACTTCTCTGATCAGCAAGAGGACACTGACAGCTGTACAGAGTCTGTGAAGTTCTG 1140
1255 GACCTTAGTCACTTGGAAAGATGTGCGAGCGATCGCTCACACCGAGCTGGTAGATGCAA 1314
1141 ATGCTCGCTCAGTGACAGCTTTGCTTCTCCCACTCTCCCACTTCAAAAATGGGCCCACTCTCAAG 1200
1315 GAAGAGAGCTTTGGGTGTACCTTCTGCTCTGAGAAATGGCAACCTTTAAAGCTGCGATGGCC 1374

QY 1201 AGAGGATGAAGTCTTATAAAACCTGCACTGATCACCCTTTATCTCATTTGTGTTCTAGTTC 1260
Db 1375 GCGAGACGGCACCTT-TAACCGAGACCTCATCACCCAGGTTAAGATCAAGGTCTT-TTC 1431
QY 1261 TCFTGCCCATCATTTGGCATAGTGGCAGCTCAGCTCTCTGAAATGGGAAACGAAGAAATTGCA 1320
Db 1432 ACCTGGCCGCA TGGACACCCAGACAGGTCCCCTACATCGTGACCTGGGAAGCTTTGGC 1491
QY 1321 CGTTGGCTCAGTTAATGACAGATATATCTCAAGTCCGGAGGCAAAAGAAATGGCAGTG 1380
Db 1492 TTTTGAACCCCTCCCTGGGTCAAGCCCT-----TTGTA CACCCCTAAGCCCTCCGCTCCT 1546
QY 1381 AAGATGAATAGATTTCGAGAAAGCTGTGTATGGAACGATGAGCAACATGAAAGACAGAA 1440
Db 1547 CTTCTCCATCCGCCCGCTCTCTCCCTTTGAACCTCTCTGTTGACCCCGCTCGATCC 1606
QY 1441 TCAGATATCTTTT CAGATAATGAAGCCAAATCTCTAGATGCTAAGAAATTTCCAAAATTTCA 1500
Db 1607 TCCCTTTATCCAGCCCTCACTCTCTCTAGGCGCCGGAATTCGTTAACTCGACATGGAA 1666
QY 1501 GCATAACACTGATCAAGATTTAATGATGTTCTTTTCCAGCTAAATTTCTTACTTTCTCT 1560
Db 1667 GTCTTTCTCTCTCGCAGCTGTGGGCTTTGCTGGGCACAGTACAATCCCAACACTCAG 1726
QY 1561 CCATCCAGG--AACATGAGAAATATCATAGGGGATATCTCCAAGTCAATTAGTAGTCTGAA 1618
Db 1727 GCTGGGAGGACATCTATCGTGCATCTCTTTGAATGGCGCTGGGCCGACATTTGCACTGGAG 1786
QY 1619 CACCACAGTACTTGATTGAGTTTCAGTATTTGAAACTGAATGGCAGAGTCCAAAGAA 1678
Db 1787 TCGCAACACTATTTAGCTCTCTAATGGTTTGGAGGAGTTTCAGGTTTCTCTCCCAAATGAA 1846
QY 1679 TGCAATTAACAACAGAGAGAGATGCGTAAATTTAGAGGCGGTATATACAATGATCAGC 1738
Db 1847 AACATTTGCTATTACTAATCCGAAACAGCCCTTGGTGGGAAAGATACCAAGCCCATCAGCTAC 1906
QY 1739 AGAAATTTAAGTCTCTAGATGAAACAAAGTATATTTTGGAAACAGGAAATAAAGGGGAAAT 1798
Db 1907 AAGATCTGAGTCGATCGGGCAATGAAATGAATTCAGAGACATGGTGACAGATGCAAC 1966
QY 1799 GAAACTGTTGAATAATATCACTAATGATCTGAGGCTGAAGGATTTGGGAACATTTCTCAGAC 1858
Db 1967 AATGTTGGAG-TTCGTAATTTATGTTGATGCTGTTGTCAATCACATGTGTGATCTATGGG 2025
QY 1859 ATTGAAAAATATCACTTTACTCCAAGTGCAGAAAGTGTCTCGTGACTGGGAAATGGAC 1918
Db 2026 TGGCACGGGCACCCACTCAACATGTGGGAGCTATTTCAACACCCGGGACTAGAGATTTTCC 2085
QY 1919 CAACGATCTGGGCTCCAAACATGACCAATCGGGGCTGTGAAACAGACAGAGGTGAATTCACAGG 1978
Db 2086 CGCTGTGCGTACTCTGCTCGGGATTTCAATGACGGCAATGTACACTGCAAGTGAGGA 2145
QY 1979 CACCTACATCACAGCCGTAAACAGCCACATCAAAATGAGATCAAAAGAGTCAACCATGCAATGG 2038
Db 2146 CATCGAAATTTATGGGACATGTATCAGGTCGCGGATTTGCAAGTTGTTCAGGCTTCTTGA 2205
QY 2039 GACA-CAAAAACCATCAACAGAGGACCCAGCCACCTTTGGCTTCACCGTCAATTTGGA 2097
Db 2206 TCTGGCTCTGAGAAAGACTATGTACGCTCAAAATTTGACGCGTACATGAATCACTCAT 2265
QY 2098 AGTTTTCAGAGTCCACCACTGTTCTTCCAGGCGCAGTCTTCATAGACAGGAATGGGAAGG 2157
Db 2266 TGATA TGGGTGAGAGGGTTCCGGATCGATGCTGCCAAGCATATGTGGCCAGGGGACAT 2325
QY 2158 AGGTCTCTGAAGACCATGTGGCTGCTGGGTCAAGTGTTAATGACATTTGGTGATGACTGGA 2217
Db 2326 AAGAGCGTTTCTGGAACAACTGCACGATCTAAATACTCACTAGTGGTTTTCAGCAGGAACGAA 2385
QY 2218 AAGCTACAGGGTC-GGCATCAACATCTTCACTCGCTCGGCACACAGAAAGGATGAGTG 2276
Db 2386 ACCCTTTATTTTACCAAGAGGTAATTTGACTTTGGGAGGAGAGCCAAATCAACAGCGATCAGTA 2445

QY 2277 AGTGACCAAGTCTCTCTG---GACTCCAGGTGAAAGAGAGATAGAGGCCCTCTCTGGAC 2333
Db 2446 CTTTGGGAATGGCCGCGTGAAGAAATTCAGATATGTGTCGCAAACTGGGGAGCGGTGATCCG 2505
QY 2334 AAAATGGTATACCAAGGCTTTCCAGGTCTAAATAGGTACTCCAGGTCTTAAAGGTGATCGGG 2393
Db 2506 GAAGTGAATGGAGAGAGATGGCTACTTAAAGAACTGGGGAGAGGCTGGGCTTTGT 2565
QY 2394 GGGATCTCTGTTTACCTGGAGTTCGAGGATTCAGGAC--CAATGGGGAAGACCGGGA 2451
Db 2566 GCCTTCTGACAGAGCCCTGGTGTGTGTGTAATACCAACGACACCAAGCGGGGGCACGGGGC 2625
QY 2452 AGCCAGGACTTAATGACAAAGGCCAGAGGAGGAAAGAGGAGTCGAGCATGCNAA 2511
Db 2626 AGGCGAGCTTCCATTTCTTCTGGAATGCCAGGCTTTATAAATGGCGGTGGTTT 2685
QY 2512 GACAATCTAATACAGTCCGACTGGTGGGTGGGCGGCCCTCAGCAAGGACAGATGGAGA 2571
Db 2686 CATGCTCGCTCATCGTACGGGTTTACACGGGTGATGTCAAGTTATCTGTTGGCCAAAGTA 2745
QY 2572 TTTTTCAGAAAGGCGATGGGGTACGGTGTGTGACGACCGCTGGGAACCTGGGTGGAGAC 2631
Db 2746 TTTCCAAAACGGAGTGGATGTTAAACGACTGGGTGGGACCAACCAAGTAACCTGGACGGATC 2805
QY 2632 T----GGTCTGTCGAGGAGCTTGGGATACAAAGGTGTTCAAAGTGTGCATAAGCGAGCT 2687
Db 2806 GACGAAGTCCGTTTCAATCAATCAACGAGACACTACCTGTGGCAATGACTGGGTCTCGAACA 2865
QY 2688 TATTTTGGAAAGGTACGGGTCCCAATATGGCTGAATGAATGATTTTGGTTCGGGAAAGAG 2747
Db 2866 TCGCTGGGCAAAATAGG-----AACATGGTTATCTTCCGTAATGTGGTAGCGGTGAGC 2921
QY 2748 TCATCCATTAAGAGTGCAGAAATPAGACAGTGGGTGTGAGAGGCTGTGTCGACGACGAA 2807
Db 2922 CTTTCTCAAACTGGTGGGACAAAGGAGCAATCAAGTAGCTTTCCGTGGCGGACACAGAG 2981
QY 2808 GATGCTGGGGTCACTTGGACCTACATATGATGATATTTTCAATTCATATTTTAA 2867
Db 2982 GCTTCATTTGCTTTAATAATGATGACTGGTATATGAATGCTGATTTGCAAACTGGTCTGC 3041
QY 2868 CTGTTTAAAGTGAATTTTTCCTTCTGCTTCACTAAATCAGCTTAATTAATTTAAGA 2927
Db 3042 CTGCTGNAACCTACTCGGATGTTATTTCTGACAAAGGAGGAGCGGTGTTACTGGAA 3101
QY 2928 AACTAAGAAATTTTATCCAGAAAAGGAATATTTAAATACTCTGATAAATCAATATAA 2987
Db 3102 AGCAGGTGTACGTTTCTCGGAGTGAAGGCCAATTTCCAGATTAGTAAACAGCGATGAAG 3161
QY 2988 TAGCTTCATATTTGCTTCAATACAGAACCATTTCACTTCTCTAGGTTTTTAAGTGGC 3047
Db 3162 ATCCATTTGTTGCAATTCACGTTGATGCCAAGTTTATAAGCTTCCAGGATCCACTAGTAAC 3221
QY 3048 TCGTGCCGAATGATCCCTCAGGATATAGTATGTTTTCGCTTTTCATAGGAGGGGAAA 3107
Db 3222 GGCGCCAGTGTGCTGGAATTCGG-----CTTGTGCACTATAGGGCGGCCAATTCGCC 3276
QY 3108 TGTAAGTCTTATGCAATCTCTGTAGTCTTTCGAACATGTTAAACCATGATGATAGCAATG 3167
Db 3277 CCTCTCCCCCCCCCTAAGTTACTGGCGAAGCGCTTGGATTAAGGCCGTGTGTG 3336
QY 3168 CTTTAAAGGAGAGAAAAGCAGCGTGATCCGATTTGGTGGAAAGTAAAGTGTGATCATC 3227
Db 3337 TTTGTCTATATGTTGATTTTCCACCATATTTGCCATATTTGCCA-----TGTGAGGGCC 3390
QY 3228 GTGCCCTTATAGGAAGGCACAGAGGGTCTGACATGATGATGATGGAACCACTGAATTC 3287
Db 3391 CGGAAACCTGGCCCTGTCTTTCGACGAGCATTCCTAGGGGTCTTTCCCTCTCGCCAAA 3450
QY 3288 GCAATTCAGAGATATTTGATTAATGATGCTAGCTCGATACAGCAAAACGCCATTTGACCAT 3347
Db 3451 GGAATGCAGAGTCTGTTGAATGTCGTGAGGAGCAGTTCCTCTGGAAGCTTCTTGAAGA 3510
QY 3348 TCACCAATTGGTGTGACCTTCAAGCTTTCAGCTGCCGCAAGCACTCAGGGCGCAAGGG 3407

Db 3511 CAACAAACGTCTGTAGCAACCTTTTGAGGAGCGGAACCCCCACCTCTGGCGACAGGTGC 3570
QY 3408 CTGCTAAAGGAAGCGGAACACACGTAGAAAGCGAGTCCGACAGAAACGCTGCTGACCCCGAT 3467
Db 3571 CTCTGGCGCCAAAAGCCACCGTGTATTAAGATACACCTGCAAAAGCGGCACACCCAGTGC 3630
QY 3468 GAATGTCACTACTGGGCTATCTGGACA-AGGNAACGCAAGCGCAAGGAAGACAGG 3526
Db 3631 CAGTTGTGATTTGGATAGTTGTGGAAAGAGTCAATGGCTCTCTCAAGCGTATGTAAC 3690
QY 3527 TAGCTTCAGTGGGCTTTACATGGCATAGTCTAGACTGGCGGTTTTTATGGACCAAGCG 3586
Db 3691 AAGGGCTGAAGGATGCCAGAGGTACCCCATTTGATTTGGATCTGATCTGGGGCCTCG 3750
QY 3587 AACCGGAATTTCCAGCTGGGGCGCCCTCTGCTAAGGTGGGAAGCCCTGCAAAAGTAAACT 3646
Db 3751 GTGCACATGCTTTACATGTTGTAGTCAGGTTTAAAAAAGCTCTAGGCCCCCGAACAC 3810
QY 3647 GGATGGCTTTCTGGCCCAAGGATCTGATGGCGCAGGGATCAAGATCTGATCAAGAGA 3706
Db 3811 GGGACGCTGGTTTTCTTTGAAAAACAGATGATTAAGCTTTGCCACAACCCCGGATAATTC 3870
QY 3707 CAGATGAGGATCGTTTCGCATGATTCAAACAAGATGATTCGACGAGGTTCTCGGCCG 3766
Db 3871 CTGCAGCAATATGGGATCGGCCATTTGAACAAGATGATTTGACGCAAGTTCTCGGCCG 3930
QY 3767 CTTGGGTGGAGAGGCTATTCCGGCTATGACTGGGCAACAAGCAATCGGCTGCTCTGATG 3826
Db 3931 CTTGGGTGGAGAGGCTATTCCGGCTATGACTGGGCAACAAGCAATCGGCTGCTCTGATG 3990
QY 3827 CCSCCGGTTTCCGGCTGTACGCCAGCGGCGCCCGTTCTTTTGTCAAGACCACTGT 3886
Db 3991 CCGCGGTTTCCGGCTGTACGCGCAGGGGCGCCGGTTCTTTTGTCAAGACCACTGT 4050
QY 3887 CCGGTGCTTGAATGAATGAACTGCAAGAGGAGCGCGGCTATCGTGGCTGGCCACGAGCG 3946
Db 4051 CCGGTGCTTGAATGAATGAACTGCAAGAGGAGCGCGGCTATCGTGGCTGGCCACGAGCG 4110
QY 3947 GCGTTCCTTCGCGAGCTGTCTCGACGTGCTCACTGAAGCGGGAAGGACTGGCTGCTAT 4006
Db 4111 GCGTTCCTTCGCGAGCTGTCTCGACGTGCTCACTGAAGCGGGAAGGACTGGCTGCTAT 4170
QY 4007 TGGCGAAGTCCCGGCGCAGGATCTCTGTCTCATCTCACTTGTCTCTGCCGAGAAAGTAT 4066
Db 4171 TGGCGAAGTCCCGGCGCAGGATCTCTGTCTCATCTCACTTGTCTCTGCCGAGAAAGTAT 4230
QY 4067 CCAATCATGCTGATGCAATCGCGGCGCTGATAGCGTTGATCGGCTACCTGCGCCATTCG 4126
Db 4231 CCAATCATGCTGATGCAATCGCGGCGCTGATAGCGTTGATCGGCTACCTGCGCCATTCG 4290
QY 4127 ACCAACAGGAAACATCGCATCGAGGAGCAGTACTCGGATGGAAGCGGCTCTTGTGCG 4186
Db 4291 ACCAACAGGAAACATCGCATCGAGGAGCAGTACTCGGATGGAAGCGGCTCTTGTGCG 4350
QY 4187 ATCAGATGATCTGGAGAAAGACATCAGGGGCTCGCGCCAGCGCAACTGTTGCCAGGC 4246
Db 4351 ATCAGATGATCTGGAGAAAGACATCAGGGGCTCGCGCCAGCGCAACTGTTGCCAGGC 4410
QY 4247 TCAAGGGCGCATGCCGACGCGGAGGATCTGCTGTGACCCATGCGGATGCTGCTTGC 4306
Db 4411 TCAAGGGCGCATGCCGACGCGGAGTATCTGCTGTGACCCATGCGGATGCGCTTGTGCG 4470
QY 4307 CGAATATCATGTTGGAATAATGGCGGCTTTTCTGATTCATCGATGTCGCGGCTGGGTG 4366
Db 4471 CGAATATCATGTTGGAATAATGGCGGCTTTTCTGATTCATCGATGTCGCGGCTGGGTG 4530
QY 4367 TGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGCTGATATTTGCTGAAGAGCTTGGCG 4426
Db 4531 TGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGCTGATATTTGCTGAAGAGCTTGGCG 4590
QY 4427 GCBAATGGGCTGACCGCTTCTCTGCTTTTACGGTATTCGCGCTCCGATTCGAGCGCA 4486

Db 4591 GCGAATGGGCTGACCGCTTCTCGTGTACGGTATCGCGCTCCCGATTCGCAGGCGCA 4650
Qy TCAGCTTCTATCGCTTCTTCCAGCAGTCTTCTGAGCGGAGCTCTGGGTTTCGATATAAT 4546
Db TCGCTTCTATCGCTTCTTCCAGCAGTCTTCTGAGCGGAGCTCTGGGTTTCGATATAAT 4710
Qy 4547 AAAAGATTTTATTTAGTCTCCAGAAAAAGGGGGAATGAAGAGACCCACCTGTAGTTTG 4606
Db 4711 AAAAGATTTTATTTAGTCTCCAGAAAAAGGGGGAATGAAGAGACCCACCTGTAGTTTG 4770
Qy 4607 GCAAGCTAGCTTAAGTAAAGCCTATTTGCAAGGCAATGAAGAAAAATACATACTGAGATAG 4666
Db 4771 GCAAGCTAGCTTAAGTAAAGCCTATTTGCAAGGCAATGAAGAAAAATACATACTGAGATAG 4830
Qy 4667 AGAAGTTCAGATCAAGGTCAGGAACAGATGAAACAGCTGATATATGGGCCAACAAGGATAT 4726
Db 4831 AGAAGTTCAGATCAAGGTCAGGAACAGATGAAACAGCTGATATATGGGCCAACAAGGATAT 4890
Qy 4727 CTGTGGTAAGCAGTTCTGTCGCCCGCTCAGGGCCAAAGAACAGATGGAACAGCTGAATATG 4786
Db 4891 CTGTGGTAAGCAGTTCTGTCGCCCGCTCAGGGCCAAAGAACAGATGGAACAGCTGAATATG 4950
Qy 4787 GGCACAAACAGATATCTGTGTAGAGCAGTTCTGTCGCCCGCTCAGGGCCAAAGAACAGATG 4846
Db 4951 GGCACAAACAGATATCTGTGTAGAGCAGTTCTGTCGCCCGCTCAGGGCCAAAGAACAGATG 5010
Qy 4847 GTCCCCAGATCGGGTCAGGCCCTCAGCAGTTCTTAGAGAACCATCAGATGTTTCCAGGT 4906
Db 5011 GTCCCCAGATCGGGTCAGGCCCTCAGCAGTTCTTAGAGAACCATCAGATGTTTCCAGGT 5070
Qy 4907 CCCCCAAGGACCTGAATAGACCTGTGCTTATTTGAACATAACCAATCAGTTCCGTTCTC 4966
Db 5071 GCCCCAAGGACCTGAATAGACCTGTGCTTATTTGAACATAACCAATCAGTTCCGTTCTC 5130
Qy 4967 GCTTCTGTTCGCGCTTCTGCTCCCGAGCTCAATAAAGAGCCCAACACCTCTCACTC 5026
Db 5131 GCTTCTGTTCGCGCTTCTGCTCCCGAGCTCAATAAAGAGCCCAACACCTCTCACTC 5190
Qy 5027 GGGGGCCAGTCCCTCCGATTGACTGAGTCGCGCGGTACCGGTATCCCATTAACCTC 5086
Db 5191 GGGGGCCAGTCCCTCCGATTGACTGAGTCGCGCGGTACCGGTATCCCATTAACCTC 5250
Qy 5087 TTGCAGTTGCACTCGACTTGTGGTCTCGCTGTCTTGGGAGGCTCTCCTCTGAGTGATT 5146
Db 5251 TTGCAGTTGCACTCGACTTGTGGTCTCGCTGTCTTGGGAGGCTCTCCTCTGAGTGATT 5310
Qy 5147 GACTACCGCTCAGCGGGGTCTTTCAATTTGG 5177
Db 5311 GACTACCGCTCAGCGGGGTCTTTCAATTTGG 5341

RESULT 5
AR381629
LOCUS AR381629 7699 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 1 from patent US 6608238.
ACCESSION AR381629
VERSION AR381629.1 GI:40089782
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 7699)
Gavara,J.S.; Falconer,M.M.; Nguyen,T.H. and Benkel,B.P.
Trans-somatics with gene transfer into mammary epithelial cells
Patent: US 6608238-A 1 19-AUG-2003;
JOURNAL Gala Design Inc.; Middleton, WI;
CAX;
FEATURES
source Location/Qualifiers
1..7699
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN

Query Match 40.2%; Score 2078.8; DB 6; Length 7699;
Best Local Similarity 64.3%; Pred. No. 0;
Matches 3339; Conservative 0; Mismatches 1814; Indels 38; Gaps 14;
Qy 1 TTTGAAAGACCCACCCGTAGGTGGCAAGCTTAGCTTAAGTAAACGCCACTTTTGCAGGCAT 60
Db 175 TTTGAAAGACCCACCCGTAGGTGGCAAGCTTAGCTTAAGTAAACGCCACTTTTGCAGGCAT 234
Qy 61 GGAATAATACATTAACCTAGAAATAGAAAAGTTTCAGATCAAGSTCAGGAACAAAGAAACAGC 120
Db 235 GGAATAATACATTAACCTAGAAATAGAAAAGTTTCAGATCAAGSTCAGGAACAAAGAAACAGC 294
Qy 121 TGAATACCAACAGGATATCTGTGGTAAAGCGTTCTGCCCCCGCTCAGGGCCCAAGAACAC 180
Db 295 TGAATACCAACAGGATATCTGTGGTAAAGCGTTCTGCCCCCGCTCAGGGCCCAAGAACAC 354
Qy 181 GATGAGACAGCTGAGTGATGGGCCAAACAGGATATCTGTGGTAAAGCTTCTCTGCCCGG 240
Db 355 GATGAGACAGCTGAGTGATGGGCCAAACAGGATATCTGTGGTAAAGCTTCTCTGCCCGG 414
Qy 241 CTGGGGCCCAAGAACAGATGGTCCCCAGATGCGGTCCAGCCCTCAGCAGTTTCTAGTGAA 300
Db 415 CTGGGGCCCAAGAACAGATGGTCCCCAGATGCGGTCCAGCCCTCAGCAGTTTCTAGTGAA 474
Qy 301 TCATCAGATGTTTCCAGGGTCCCCAAGAGACCTGAAAATGACCCCTGTACCTTATTTGAAC 360
Db 475 TCATCAGATGTTTCCAGGGTCCCCAAGAGACCTGAAAATGACCCCTGTACCTTATTTGAAC 534
Qy 361 TAACCAATCAGTTCTCGCTTCTGTTTGGCGGCTTCCGCTCTCCGAGCTCAATAAA 420
Db 535 TAACCAATCAGTTCTCGCTTCTGTTTGGCGGCTTCCGCTCTCCGAGCTCAATAAA 594
Qy 421 AGAGCCACAAACCCCTCAGTCCGGCGCCAGTCTTCCGATAGACTGGTCCGCCGGGTAC 480
Db 595 AGAGCCACAAACCCCTCAGTCCGGCGCCAGTCTTCCGATAGACTGGTCCGCCGGGTAC 654
Qy 481 CCGTATTTCCCAATAAAGCCTCTTGTCTTTTGCATCCGAATCGTGGTCTCGCTGTTCCTTG 540
Db 655 CCGTATTTCCCAATAAAGCCTCTTGTCTTTTGCATCCGAATCGTGGTCTCGCTGTTCCTTG 714
Qy 541 GAGGGTCTCTCTGAGTGATGACTACCACGACGGGGGTCTTTTCAATTTGGGGGTCTGT 600
Db 715 GAGGGTCTCTCTGAGTGATGACTACCACGACGGGGGTCTTTTCAATTTGGGGGTCTGT 774
Qy 601 CCGGATTTGAGACCCCTGCCCCAGGACCAACGACCCACGACGGGAGGTAAAGTGGCC 660
Db 775 CCGGATTTGAGACCCCTGCCCCAGGACCAACGACCCACGACGGGAGGTAAAGTGGCC 834
Qy 661 AGCAACTTATCTGTCTGTCCGATGTTCTAGTGTCTATGTTTGTATGCGCCCTGCG 720
Db 835 AGCAACTTATCTGTCTGTCCGATGTTCTAGTGTCTATGTTTGTATGCGCCCTGCG 894
Qy 721 TCTGTACTAGTTAGCTTAAGCTCTGTATCTGCGGACCCGCTGGTGGAACTGACGAGTT 780
Db 895 TCTGTACTAGTTAGCTTAAGCTCTGTATCTGCGGACCCGCTGGTGGAACTGACGAGTT 954
Qy 781 CTGAACACCCGGCGCAACCTCGGAGACGCTCCAGGACCTTTGGGGCCCGTTTGTGG 840
Db 955 CTGAACACCCGGCGCAACCTCGGAGACGCTCCAGGACGCTTTGGGGCCCGTTTGTGG 1014
Qy 841 CCGACCTGAGGAGGAGTTCGATGTTGAATCCGACCCCGTCAGGATATGTTGTTCTGGT 900
Db 1015 CCGACCTGAGGAGGAGTTCGATGTTGAATCCGACCCCGTCAGGATATGTTGTTCTGGT 1074
Qy 901 AGGAGACGAGAACCTAAACAGATTCCGCTCCGCTCTGAAATTTTGTCTTCGGTTTGAA 960
Db 1075 AGGAGACGAGAACCTAAACAGATTCCGCTCCGCTCTGAAATTTTGTCTTCGGTTTGAA 1134
Qy 961 CCGAGCCGCGCTCTTGTCTGTGAGCCAGCTTGGGCTGCGAGTTCGACTCTAGAGGA 1020
Db 1135 CCGAGCCGCGCTCTTGTCTGTGAGCCGCTGCGAGCTTCGTTCTGTCTGTCTCTGTCT 1194
Qy 1021 TCAATTCGACACGAGTAAATCGGTGCTGCGGTCTTTAGGACATATGAAGTATGGCACAGT 1080

| | | | |
|----|------|--|------|
| Db | 1195 | GACTGTGTTCTGTATTTGTTCTGAAATATTAGGGCCAGACTGTTACCACTCCCTTAAGTTT | 1254 |
| Qy | 1081 | GGGATGACTTTTCTTGATCAGCAAGAGGACACTGACAGCTGTACAGAGCTGTGTGAAGTTGC | 1140 |
| Db | 1255 | GACCTTAGGTCACITGGAAAGATGTCGAGCGGATCGCTCA CNA CCAAGTCGGTAGATGTCAA | 1314 |
| Qy | 1141 | ATGCTCGCTCAGTGAAGCTTTGTTCTCTCCCATCCTAAATAATGGGCCCAACTCTTCAAG | 1200 |
| Db | 1315 | GAAGAGAGTGTGGGTTACCTTCTGCTCTGAGAAATGGCCAACTTTAACTCGGATGGCC | 1374 |
| Qy | 1201 | AGAGGATGAGTCTTATATAAAGTGCATGATCAACCTTTATCTCATTTGTGTGTGTAGTTC | 1260 |
| Db | 1375 | GCGAGACGGCACCTT-TAA CCGAGACCTCATCACCCAGGTTAAGATCAAGGTCTT--TTC | 1431 |
| Qy | 1261 | TCGTGCCCATCATTTGGCATAGTGGCAGCTCAGCTCTGAAATGGAAATGGAACGAAGAATTGCA | 1320 |
| Db | 1432 | ACCTGGCCCGGATGGACACCCAGACAGGTCCTCTACATCGTACCTGGGAGAGCCTTGGC | 1491 |
| Qy | 1321 | CGGTTGGCTCAGTTAAATGCAAGATATATCTCCAAAGTCGGAAGCAAGAAATGGCAGTG | 1380 |
| Db | 1492 | TTTTGACCCCTCTCCCTGGGTCAAGCCCT-----TTGTACACCTAAGCCTCGCCCTCT | 1546 |
| Qy | 1381 | AAGATGAAATGAGATTTTCGAGAGCTGTGATGGAAACCATGAGCAACATGGAAGACGAA | 1440 |
| Db | 1547 | CTTCCTCCATCCGCCCCGTCTCTCCCTTTGAACCTCTCTGTTGCAGCCCGCTCGATCC | 1606 |
| Qy | 1441 | TCCAGTATCTTTACAGATAATGAAGCCAAATCTCTTAGATGCTAGAAATTTCCAAAATTTCA | 1500 |
| Db | 1607 | TCCCTTTATCCAGCCCTCATCTCTTCTTAGGCGCCGGAAATTCGTTAACTCGCATGGAA | 1666 |
| Qy | 1501 | GCATAACAACCTGATCAAGAATTTAATGATGTTCTTTTCCAGCTAAATTCCTTACTTTCTCT | 1560 |
| Db | 1667 | GTCTTCTCTCTCTCGCAGCTGTGCGGCTTTGCTGGGCACAGTACAAATCCCAACACTCAG | 1726 |
| Qy | 1561 | CCATCCAGG--AACATGAGAATATCATAGGGGATATCTCCAGTCAATTAGTAGGTCTGAA | 1618 |
| Db | 1727 | GCTGGGAGGACATCTATCTGTGCATCTCTTTGAAATGGCGCTGGGCCGACATTCACACTGGAG | 1786 |
| Qy | 1619 | CACCACAGTACTTGATTTGTCAGTTTCAGTATTTGAAACATCAATGGCAGAGTCCCAAGGAA | 1678 |
| Db | 1787 | TGGCAACATATTAGTCTCTATTTGGTTTGGAGGAGTTCAAGTTTCTCTCCCAATGAA | 1846 |
| Qy | 1679 | TGCATTTAAACAACAAGGAGATGCGTAAATATTAGAGGAGCGTATATCAATGCATCAGC | 1738 |
| Db | 1847 | AACATTGTCATTACTAATCCGAACAGCCCTGGTGGGAAAGATACCAGGCCATCAGCTAC | 1906 |
| Qy | 1739 | AGAAATTAAGTCTCTAGATGAAACCAAGTATATTTGGAAACAGGAATAAAGGGGAAAT | 1798 |
| Db | 1907 | AAGATCTGCAGTCGATCGGCGCAATGAAATGAATTCAGAGACATGTTGACAGATGCAAC | 1966 |
| Qy | 1799 | GAAACTGTTGAAATATATCACTAATGATCTGAGGCTGAAGGATGGGAAACATTTCTCAGAC | 1858 |
| Db | 1967 | AATGTTGGAG--TTGCTATTTATGTGAGTCTGTTGTCAATCAATGATGTGTGATCTATGGG | 2025 |
| Qy | 1859 | ATTGAAAAATATCACTTTACTCCAAAGTGCAGAAAGTGTCTGCTGACTGGGAAATGGAC | 1918 |
| Db | 2026 | TGGCACGGGCACCCACTCAACATGTGGAGGTATTTCAA CACCGGACTAGAGATTTTC | 2085 |
| Qy | 1919 | CAACGATCTGGGCTCCAAATGACCATCGGGGCTGTGAAACAGCAGAGGTGAATTCACAGG | 1978 |
| Db | 2086 | CGCTGTCCGCTACTCTGCTGGGATTTCAATGACGGCAATGTGCACACTGCAAGTGGAGA | 2145 |
| Qy | 1979 | CACCTACATCACAGCCGTAACAGCCACATCAATATGAGATCAAGAGTCCACCATGCGATGG | 2038 |
| Db | 2146 | CATCGAAAAATTATGGGACATGTATCAGGTCGCGGATTTGCAAGTTGTCAGCCTCTCTGA | 2205 |
| Qy | 2039 | GACA-CAAAAAACCACTCAAACAAGAGGACCAGACCCACCTTTTGGCTTCAACGTCAAATGGA | 2097 |
| Db | 2206 | TCTGGCTCTGGAGAAAGACTATGTACGCTCAACAAATTGACGGGTACATGAATCACCTCAT | 2265 |
| Qy | 2098 | AGTTTTCAGAGTCCACCACTGTCTTCA CCGGCCCATGCTTTCATAGACAGGAATGGGAAGG | 2157 |

| | | | |
|----|------|--|------|
| Db | 2266 | TGATATGGGTGTAGCAGGGTTCCGGATCGATGCTGCCAAGCATATGTGGCCACGGGACAT | 2322 |
| Qy | 2158 | AGGTCTCGAAGACCATGTGGCTGCTCGGGTCAAAGTGTTAATGACATTTGGTGATGACTGGA | 2217 |
| Db | 2326 | AAGAGCGTTTCTGGACAAACTGCACGATCTAATACTCAGTGGTTTTACAGGACGAA | 2385 |
| Qy | 2218 | AAGCTACAGGGTC - GGCACTCAACATCTTCACTCGCCCTGGGCACACAGAAAGNGTAGTGG | 2276 |
| Db | 2386 | ACCCTTTATTTTACCAAGAGGTAAATTGACTTGGGAGGAGAGCCAAATCACAGGCAGTCA | 2445 |
| Qy | 2277 | AGTGACCAAGGTCTCTCTG - -GACTCCAGGTGAAAAAGGAGATAGAGGCCCTCTCTGGAC | 2333 |
| Db | 2446 | CTTTGGGAATGGCCCGCTGACAGAAATCAAGTATGTGTGCCAACTGGGACCGGTATCCG | 2505 |
| Qy | 2334 | AAAATGCTATACAGGCTTTTCCAGGCTTAATAGGTACTCCAGGTCTTTAAAGGTGATCGGG | 2393 |
| Db | 2506 | GAAGTGGAAATGGAGAGAAGATGCGCTACTTTAAAGAACTGGGAGAAAGCTGGGGCTTTGT | 2565 |
| Qy | 2394 | GGGATCTCTGGTTTACTCTGAGTTTCAGAGATTCCTCAGGAC - -CAATGGGGAAGACCGGA | 2455 |
| Db | 2566 | GCCTTCTGACAGAGCCCTGCTGTGTGTGTGATATACCAACGCAACACGCGGGGCA | 2625 |
| Qy | 2452 | AGCCAGGACTTAATGGACAAAAGGCCAGAGGGAGAAAAGGAGTGGAAACATGCARA | 2511 |
| Db | 2626 | AGGCGGAGCTTCCATCTTACTTCTGGGATGCGCAGGCTTTATAAATGGCGGTGGTTTT | 2685 |
| Qy | 2512 | GACAACTTAATACAGTCCGACTGGTGGTGGACGCGGCCCTCACGAAGCAGAGTGGAGA | 2571 |
| Db | 2686 | CATGCTCGCTCATCCGTACGGGTTTCAACGGGTGATGTCAAGTTATCTGTGGCCAAGATA | 2745 |
| Qy | 2572 | TTTTTCAAGAGCGACGTGGGGTACGGTGTGTGACGACCGCTGGGAACTCGGTGGAGGAC | 2631 |
| Db | 2746 | TTTTCGAAAAAGGAGTGGATGTTAAACGACTGGGTGGGACCAACCAAGTAACTCGACGCGATC | 2805 |
| Qy | 2632 | T - - - -GGTCTCTCGAGGACTTGGGATCAAAAGGTGTTCAAGTGTGCATAAGCCAGCT | 2687 |
| Db | 2806 | GACGAATCCGTTACAATCAACGACGACACTCTGTGGCAATGACTGGGTCTCGCAACA | 2865 |
| Qy | 2688 | TATTTTCGAAAAAGGTACGGGTCCAAATATGCTGAAATGAAAGTATTTTGTTCGGGAAAGAG | 2747 |
| Db | 2866 | TCGCTGGCGCAAAATAGG - - - -AACATGGTTATCTTCGTAATGTGTAGACGGTCA | 2921 |
| Qy | 2748 | TCAATCCATTGAAGAGTCAGAAATTAGACAGTGGGGTGTGAGAGCCCTGTTTCGACGACGAA | 2807 |
| Db | 2922 | CTTTCTCAAACTGTTGGGACAAACGGGAGCAATCAAGTAGCTTTCGGTTCGCGGCGACAGAG | 2981 |
| Qy | 2808 | GATGCTGGGGTCACTTTGGACCTCAATAATGATCATATTTTTCATTCACATTTTTTTAAA | 2867 |
| Db | 2982 | GCCTCATTTGTCTTTTAATAATGATGACTGGGTATATGAATGTGCAATTTGCAACTGGTCTGC | 3041 |
| Qy | 2868 | CTGTTATAAGTGTATTTTTTCTTTCTTTCTTCAGTAAATTCAGCTTAATTAATTAATAAGA | 2927 |
| Db | 3042 | CTGCTGGAACCTACTCGGATGTTTATTTCTGGACAAAGGAAGGACGCTGGGTACTGGAA | 3101 |
| Qy | 2928 | AACATAAGAAATTTTATCCACAGAAAAGGAATATTTAAAAATCACTGATATAAATATAAAA | 2987 |
| Db | 3102 | AGCAGGTGTACGTTTCTTCGGATGGAAAGGCCAAATTTCCAGATTTAGTAACAGCGATGAAG | 3161 |
| Qy | 2988 | TAGCTTCATATTTGCTTCAAATACAGAACCATTTCCACTCTCTAGTGTTTTTTAAGTGGC | 3047 |
| Db | 3162 | ATCCATTTGTTGCAATTTACGTTGTATGCCCAAGTTATAAGCTTCGAGGATCCACTAGTAAC | 3221 |
| Qy | 3048 | TCGTGCCGAATTCGATCCCTTCAGGATATAGTAGTGTTCGTTTTTGCATAGGGAGGGGAAA | 3107 |
| Db | 3222 | GGCCGCCAGTGTGCTGGAAATCCGG - - - -CTTGTGCACATCTTAGGGCGGCCAATTCGCC | 3276 |
| Qy | 3108 | TGTAGTCTTATGCAATACTCTGTAGTCTTTGCAACATGGTAAGCATGAGTTTAGCAACATG | 3167 |
| Db | 3277 | CTCTCCCCCCCCCCCTTAACGTTACTGGCCGAAGCCGCTTGGAAATAGGCCGGTGTGTG | 3336 |
| Qy | 3168 | CTTTAAGGAGAGAAAAAGCAACCGTGCAATGCCGATTTGGTGGAAATGAAGGTGGTAGATC | 3227 |
| Db | 3337 | TTTGTCTATATGTGATTTTCCACATATTCGCGCTTTTGGCAA - - - -TGTGAGGGCC | 3390 |


```
Qy 3228 GTGCTTATTAGGAAGCAACAGACGGCTGTGACATGGATTGGAGCAACCACTGAATTCC 3287
Db
Qy 3391 CGGAACCTGGCCCTGTCTTTCTTGACGAGCAATTCCTAGGGTCTTTCCCTCTCGCCAAA 3450
Db
Qy 3288 GCATTGACAGATATTTGATTTAAATGCTAGCTCGATACAGCAAAAGCCATTTGACCAT 3347
Db
Qy 3451 GGAATGCAAGGTCTGTTGAAATGCTGTAAGGAAGCAGTTCCTCTGGAAGCTTCTTGAAGA 3510
Db
Qy 3348 TCACCAATTTGGTGTGCACTTCAAGCTTCAAGCTGCGGCAAGCACTCAGGGCGCAAGGG 3407
Db
Qy 3511 CAAACAACGTCTGTAGCGACCTTTGCAAGGAGCGGAACCCCCACCTGCGGCAAGGTGC 3570
Db
Qy 3408 CTGCTAAAGGAAGCGGAACACAGTAGAAGCCAGTCCCGAGAAAGCGTGTGACCCCGGAT 3467
Db
Qy 3571 CTCTGGGCGCAAAAGCCACGTGTATAGATACACTGTGAAAGCGCGCAACACCCCACTGC 3630
Db
Qy 3468 GAATGTCAGCTTACTGGGCTATCTGGACA - AGGGAAAACGCAAGCGCAAGAGAAAGCAGG 3526
Db
Qy 3631 CACGTTGTGAGTTGGATAGTTGTGMAAGAGTCAATGGCTCTCTCAAGGTAAGTCAAC 3690
Db
Qy 3527 TAGCTTGACGTGGGCTTACATGGCGATAGTACTGGGCGGTTTATGACACGAAGCG 3586
Db
Qy 3691 AAGGGGCTGAAGGATGCCAGAGGTACCCCAATGTATGGGAATCTGATCTGGGGCCCTCG 3750
Db
Qy 3587 AACCGGAATTCAGCTGGGGCGCCCTCTGTGAAGTTGGGAAGCCCTGCAAAATAACT 3646
Db
Qy 3751 GTGCACATGCTTTACATGTGTTTGTAGTCAGGTTTAAAAAGCTCTAGGCCCCCGCAACCAC 3810
Db
Qy 3647 GGATGGCTTTCTTGCCCGCAAGGATCTGATGGCGCAGGGATCAAGATCTGATCAAGAGA 3706
Db
Qy 3811 GGGACGTGGTTTCTTTGMAAACAACGATGATAAGCTTGCCACAACCCGGGATAATTC 3870
Db
Qy 3707 CAGATGAGGATGCTTTGCGATGATGAAACAAGATGATGACGCAAGTTCTCCGCGCG 3766
Db
Qy 3871 CTGCAAGCAATATGGGATCGGCCATGAAACAAGATGATGACGCAAGTTCTCCGCGCG 3930
Db
Qy 3767 CTTGGGTGGAGAGCTATTCGGCTATGACTGGGCACACAGACATCGGCTGCTCTGATG 3826
Db
Qy 3931 CTTGGGTGGAGAGCTATTCGGCTATGACTGGGCACACAGACATCGGCTGCTCTGATG 3990
Db
Qy 3827 CCGCGGTGTTCCGGCTGTACGCGCAGGGCGCCCGGTTCTTTTGTCAAGACCACTGT 3886
Db
Qy 3991 CCGCGGTGTTCCGGCTGTACGCGCAGGGCGCCCGGTTCTTTTGTCAAGACCACTGT 4050
Db
Qy 3887 CCGGTGCCCTGAATGAATGACGACGAGGCGCGGCTATCGTGGCTGCCCAACGACGG 3946
Db
Qy 4051 CCGGTGCCCTGAATGAATGACGACGAGGCGCGGCTATCGTGGCTGCCCAACGACGG 4110
Db
Qy 3947 GCGTTCTTGGCAGCTGTGCTCGACGTTGCTCACTGAAGCGGGAAGGACTGGCTGCTAT 4006
Db
Qy 4111 GCGTTCTTGGCAGCTGTGCTCGACGTTGCTCACTGAAGCGGGAAGGACTGGCTGCTAT 4170
Db
Qy 4007 TGGCGGAAGTGC CGGGCGCAGGATCTCTGTCTCATCTCACTTGTCTCTCGCGCAGAAAAGTAT 4066
Db
Qy 4171 TGGCGGAAGTGC CGGGCGCAGGATCTCTGTCTCATCTCACTTGTCTCTCGCGCAGAAAAGTAT 4230
Db
Qy 4067 CCATCATGGCTGATGCAATCGGCGGCTGACATACGCTTGATTCGGGCTACCTGCCCATTCG 4126
Db
Qy 4231 CCATCATGGCTGATGCAATCGGCGGCTGACATACGCTTGATTCGGGCTACCTGCCCATTCG 4290
Db
Qy 4127 ACCACCAAGCGAATCATGCAATCGAGCGACGACGTAATCGGATGGAAAGCGGCTTTGTCG 4186
Db
Qy 4291 ACCACCAAGCGAATCATGCAATCGAGCGACGACGTAATCGGATGGAAAGCGGCTTTGTCG 4350
Db
Qy 4187 ATCAGGATGATCTGGAAGAGCATCAGGGGCTCGCGCGCAGCGCAAGCTGTTCGCCAGGC 4246
Db
Qy 4351 ATCAGGATGATCTGGAAGAGCATCAGGGGCTCGCGCGCAGCGCAAGCTGTTCGCCAGGC 4410
Db
Qy 4247 TCAGGGCGCATGCCCGACGGCGAGGATCTGTCGTGACCCCAATGGCGATGCTGCTTGC 4306
Db
Qy 4411 TCAGGGCGCATGCCCGACGGCGAGGATCTGTCGTGACCCATGACCATGGCGATGCTGCTTGC 4470
```

```
Qy 4307 CGAATATCATGGTGGAAAATGGCGCTTTTCTGGATTTCATCGACTGTGGCGGCTGGGTG 4366
Db
Qy 4471 CGAATATCATGGTGGAAAATGGCGCTTTTCTGGATTTCATCGACTGTGGCGGCTGGGTG 4530
Db
Qy 4367 TGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTTGCTGAAGAGCTTGGCG 4426
Db
Qy 4531 TGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTTGCTGAAGAGCTTGGCG 4590
Db
Qy 4427 GCGAATGGGCTGACCGGCTCTCTGCTGCTTTTACGGTATCGGCTCCCGATTCGAGCGCA 4486
Db
Qy 4591 GCGAATGGGCTGACCGGCTCTCTGCTGCTTTTACGGTATCGGCTCCCGATTCGAGCGCA 4650
Db
Qy 4487 TCGCCCTTCTATCGCTTCTTTGACGAGTCTTCTTCTGAGCGGACTCTGGGGTTCGATAAAAT 4546
Db
Qy 4651 TCGCCCTTCTATCGCTTCTTTGACGAGTCTTCTGAGCGGACTCTGGGGTTCGATAAAAT 4710
Db
Qy 4547 AAAAGATTTTATTAGTCTCCAGAAAAGGGGGGAATGAAAGACCCCACTGTAGGTTTG 4606
Db
Qy 4711 AAAAGATTTTATTAGTCTCCAGAAAAGGGGGGAATGAAAGACCCCACTGTAGGTTTG 4770
Db
Qy 4607 GCAAGCTAGCTTAAGTAACGCCATTTTGAAGGATGGAATAATACATTAAGTGAATAG 4666
Db
Qy 4771 GCAAGCTAGCTTAAGTAACGCCATTTTGAAGGATGGAATAATACATTAAGTGAATAG 4830
Db
Qy 4667 AGAAGTTTCAATCAAGCTCAGGAACAGATGGAACAGCTGAATATGGGCCAAACAGGATAT 4726
Db
Qy 4831 AGAAGTTTCAATCAAGCTCAGGAACAGATGGAACAGCTGAATATGGGCCAAACAGGATAT 4890
Db
Qy 4727 CTGTGGTTAAGCAGTTCCTGCCCCGGCTCAGGGCCAAAGAACAGATGGAACAGCTGAATATG 4786
Db
Qy 4891 CTGTGGTTAAGCAGTTCCTGCCCCGGCTCAGGGCCAAAGAACAGATGGAATATG 4950
Db
Qy 4787 GGGCAAAACAGGATATCTGTGTAAGCAGTTCTGTCGCCCGGCTCAGGGCCAAAGAACAGATG 4846
Db
Qy 4951 GGGCAAAACAGGATATCTGTGTAAGCAGTTCTGTCGCCCGGCTCAGGGCCAAAGAACAGATG 5010
Db
Qy 4847 GTCCCCAGATGCGGTCAGCCCTCAGCAGTTTCTAGAGAACCATCAGATGTTTCCAGGGT 4906
Db
Qy 5011 GTCCCCAGATGCGGTCAGCCCTCAGCAGTTTCTAGAGAACCATCAGATGTTTCCAGGGT 5070
Db
Qy 4907 GCGCCAAAGGACCTGAAATGACCCCTGTGCTTATTTGAACCTAAACCAATCAGTTGCTTCTC 4966
Db
Qy 5071 GCGCCAAAGGACCTGAAATGACCCCTGTGCTTATTTGAACCTAAACCAATCAGTTGCTTCTC 5130
Db
Qy 4967 GCTTCTGTTGCGCGGCTTCTGCTCCCGAGCTCAATAAAGAGCCCAACCCCTCACTC 5026
Db
Qy 5131 GCTTCTGTTGCGCGGCTTCTGCTCCCGAGCTCAATAAAGAGCCCAACCCCTCACTC 5190
Db
Qy 5027 GGGGCGCCAGTCTCTCGATTGACTGAGTGGCGCGGTTACCGGTATCCAAATAACCCCTC 5086
Db
Qy 5191 GGGGCGCCAGTCTCTCGATTGACTGAGTGGCGCGGTTACCGGTATCCAAATAACCCCTC 5250
Db
Qy 5087 TTGCAGTTGCAATCCGACTTGTGCTTGTGCTTGTTCCTTTGGGAGGCTCTCTCTGAGTGATT 5146
Db
Qy 5251 TTGCAGTTGCAATCCGACTTGTGCTTGTGCTTGTTCCTTTGGGAGGCTCTCTCTGAGTGATT 5310
Db
Qy 5147 GACTACCCGTCAGCGGGGCTTTTCATTGG 5177
Db
Qy 5311 GACTACCCGTCAGCGGGGCTTTTCATTGG 5341
```

RESULT 6

```
Q879093 LOCUS 7436 bp DNA linear PAT 11-OCT-2004
DEFINITION Sequence 6 from Patent WO2004062554.
ACCESSION Q879093
VERSION Q879093.1 GI:54033470
KEYWORDS .
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1 other sequences; artificial sequences.
AUTHORS Martinez-Serrano,A., Liste,I. and Villa,A.
```


TITLE Enhancement of neuron generation and survival
JOURNAL Patent: WO 2004062554-A 6 29-JUL-2004;
NSGENE A/S (DK)

FEATURES

Location/Qualifiers
1..7436
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Viral vector"

ORIGIN

Query Match 38.1%; Score 1972.6; DB 6; Length 7436;
Best Local Similarity 65.1%; Pred. No. 0;
Matches 3371; Conservative 0; Mismatches 1529; Indels 281; Gaps 20;

Qy 1 TTTGAAAGACCCACCGTAGGTGGCAAGCTAGCTTAAGTAAACCCCACTTTGCAAGGCAT 60
Db |||||
Qy 1 TTTGAAAGACCCACCGTAGGTGGCAAGCTAGCTTAAGTAAACCCCACTTTGCAAGGCAT 60
Db |||||

Qy 61 GGAAAAATACATACTGAGTATAGAAAAGTTTCAATCAAGTCAAGGACCAAGAAACAGC 120
Db |||||

Qy 61 GGAAAAATACATACTGAGTATAGAAAAGTTTCAATCAAGTCAAGGACCAAGAAACAGC 120
Db |||||

Qy 121 TGAATACCAACAGGATATCTGTGTAAGCGGTTCTCGCCCGGCTCAGGCGCCCAAGAAC 180
Db |||||

Qy 121 TGAATACCAACAGGATATCTGTGTAAGCGGTTCTCGCCCGGCTCAGGCGCCCAAGAAC 180
Db |||||

Qy 181 GATGAGACAGCTGAGTATGAGTGGCCCAACAGGATATCTGTGTAAGCGGTTCTCGCCCGG 240
Db |||||

Qy 181 GATGAGACAGCTGAGTATGAGTGGCCCAACAGGATATCTGTGTAAGCGGTTCTCGCCCGG 240
Db |||||

Qy 241 CTCGGGCGCAAGAACAGATGTCCTGAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAGTCCG 300
Db |||||

Qy 241 CTCGGGCGCAAGAACAGATGTCCTGAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAGTCCG 300
Db |||||

Qy 301 TCATCAGATGTTCCAGGTCGCCCAAGGACCTGAAATGACCTGTACCTTATTTGAAC 360
Db |||||

Qy 301 TCATCAGATGTTCCAGGTCGCCCAAGGACCTGAAATGACCTGTACCTTATTTGAAC 360
Db |||||

Qy 361 TAACCAATCAGTTCGCTTCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
Db |||||

Qy 361 TAACCAATCAGTTCGCTTCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
Db |||||

Qy 421 AGAGCCCAACCCCTCACTCGGCGGCGAGTCTTCGAGTACAGTCTGCTGCTGCTGCTGCTGCTGCT 480
Db |||||

Qy 421 AGAGCCCAACCCCTCACTCGGCGGCGAGTCTTCGAGTACAGTCTGCTGCTGCTGCTGCTGCTGCT 480
Db |||||

Qy 481 CCGTATTTCCCAATAAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
Db |||||

Qy 481 CCGTATTTCCCAATAAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
Db |||||

Qy 541 GGAGGGTCTCTCTGAGTATGACTACCCAGGAGGGGCTTTTCAATTTGGGGGCTCGT 600
Db |||||

Qy 541 GGAGGGTCTCTCTGAGTATGACTACCCAGGAGGGGCTTTTCAATTTGGGGGCTCGT 600
Db |||||

Qy 601 CCGGGAATTGGAGACCCCTGCCCAGGACCAACCCAGGAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 660
Db |||||

Qy 601 CCGGGAATTGGAGACCCCTGCCCAGGAGCAACCCAGGAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 660
Db |||||

Qy 661 AGCAACTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
Db |||||

Qy 661 AGCAACTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
Db |||||

Qy 721 TCTGTACTAGTTAGTAACTAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Db |||||

Qy 721 TCTGTACTAGTTAGTAACTAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Db |||||

Qy 781 CTGAACACCCCGCGCAACCCCTGCGGAGACGCTCCAGGAGACTTTGGGGGCGGCTTTTCTGG 840
Db |||||

Qy 781 CTGAACACCCCGCGCAACCCCTGCGGAGACGCTCCAGGAGACTTTGGGGGCGGCTTTTCTGG 840
Db |||||

Qy 841 CCGGACTGAGGAAGGAGTGCATGTGGAATCCGACCCCGGTCAGGATATGTGGTTCGGT 900
Db |||||

Db 841 CCGGACTGAGGAAGGAGTGCATGTGGAATCCGACCCCGTCAGGATATGTGGTTCGGT 900
Qy 901 AGGAGACGAGAACTAAAAAGTTCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCG 960
Db |||||

Db 901 AGGAGACGAGAACTAAAAAGTTCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCG 960
Db |||||

Qy 961 CCGAAGCCGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
Db |||||

Db 961 CCGAAGCCGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 988
Db |||||

Qy 1021 TCAATTCGCGACGAGTAAATCGGTCGCTCTTTAGGACATATGAAGTATGCGACAGT 1080
Db |||||

Db 989 -----CGCTGACAGCTCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1028
Db |||||

Qy 1081 GGGATGATCTTCTGATCAGCAAGAGGACACTGACGCTGTAAGAGTCTGTAAGTTCG 1140
Db |||||

Db 1029 TTCTGTATTTGCTGAAAAATTAGGGCCAGACTGTTACC-----ACTCCCTTAAGTTG 1081
Db |||||

Qy 1141 ATGCTCGCTCAGTGACAGCTTGTCTCTCCCATCTTAAATGCGCCCACTCTTCAAG 1200
Db |||||

Db 1082 ACCTTAGGTCACTGGAAGATGTCGAGCGGATCGCTCACCAACAGTCCGTAAGTGTCAAG 1141
Db |||||

Qy 1201 AGAGGATGAAGTCTTATAAAAGTGCACCTGATCACCTTTTATCTCATTTGTTGTAGTTC 1260
Db |||||

Db 1142 AAGAGAGCTTGGGTTA--CTTCTGCTCTGCAAGATGGCCAACTTTAAGTCCGATGGCC 1200
Db |||||

Qy 1261 TCGTCCCATCATTTGGCATATGTCAGCTCAGCTCTGAAATGCGAAACGAGAAATTCGA 1320
Db |||||

Db 1201 GCGAGACGGCACCTTTAAACCGAGACCTCATCACCGAGTTAAGATCAAGGTCTTTTCAAC 1260
Db |||||

Qy 1321 CGGTTGGCTCAGTTAATGCGATATATCTCAAGTCCGGAAGGCAAGAAATGCGAGTG 1380
Db |||||

Db 1261 TGGCCGCTGAGGACCCAGACAGGCTCCCTTACATCTGACCT---GGGAAGCTTTGGC 1317
Db |||||

Qy 1381 AAGATGAATGAGATTTTCGAGAGCTGTGATGGAACCTGAGCAATGAGCAACATGGAAGAGCA 1440
Db |||||

Db 1318 TTTTGACCCCTCCCTCGGCTCAAGCCCTTTGTACACCTTAAGCTCCGCTCCTCTTCC 1377
Db |||||

Qy 1441 TCCAGTATCTTTTCAAGTAAATGAAAGCCAACTCTCTAGATGCTAAGAAATTTCA 1500
Db |||||

Db 1378 TCCATCCGCGCGCTCTCTCCCTTTGAACCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1437
Db |||||

Qy 1501 GCATACCACTGATCAAGATTTAAATGATGTTCTTTTCCAGCTAAATTTCTTCTTCT 1560
Db |||||

Db 1438 TTATCCAGCCCTCACTCTCTCTAGGCGCGGAAATTCGATGTACGGGCCAGATATAGC 1497
Db |||||

Qy 1561 CCATCCAGGAACATGAGAAATCATAGGGATATCTCCAAGTCAATAGTAGTGTCTGAACA 1620
Db |||||

Db 1498 GTTGACATGATTTGCTAGTATTAATGATTAATCAATTAACGGGGTCAATAGTTCATA 1557
Db |||||

Qy 1621 CCACAGTACTTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680
Db |||||

Db 1558 GCCATATATGAGTTCGCGTTACATACTTACGGTAAATGGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1617
Db |||||

Qy 1681 CATTTAAACAAAGAGGAGATGCGTAAATTAGAGGAGCGTATATACAATGCTAGCAG 1740
Db |||||

Db 1618 CCAACGACCCCGCCCAATTCAGCTCAATAATGAGTATGTTTCCCATAGTAAAGCCCAATAG 1677
Db |||||

Qy 1741 AATTAAGTCTTAGATGAAAAACAGTATATTTGGAACAGGAAATAAAGGGGAAATGA 1800
Db |||||

Db 1678 GGAATTTCCATGACGTCATGAGTGGTGGCTATTTTACCGTAACTGCCCCCTGCGCAGTAC 1737
Db |||||

Qy 1801 AACTGTTGAATAATATACATTAATGATCTGAGGCTGAGGATTTGGGAACATTTCTCAGACAT 1860
Db |||||

Db 1738 ATCAAGTATCATATGCCAAGTACGCGCCCTATTGACGTCATGACGTTAATGCGCGC 1797
Db |||||

Qy 1861 TGAATAATATCATTTTACTCAAGGTCAGGAGTCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1920
Db |||||

Db 1798 CTTGGCAT---TATGCCAGTACATGACCTTATGGGACTTTTCTTACTTTGGCAGTACATCT 1854
Db |||||

Qy 1921 ACATCTGGGCTCAACATGACCATCGGGCTGTGAACAGCAGAGGTGAATTCACAGCA 1980
Db |||||

Db 1855 ACGTATTTAGTCACTCGCTATTACCATGGTG-----ATGCGGTTTTTGGCAGTAC 1901
Db |||||

| | | | | | |
|----|------|---|---|-------------------|------|
| Qy | 1981 | CCTACATCATCAGCCGGTAA | CAGCCACATCAAAATGAGATCAAAGAGTCA | CCACTCGATGGGA | 2040 |
| Db | 1902 | ATCAATGGGGTGGATAGCGTTTGA | CTCAGCGGGATTTCAAAGTCTCCACCCCATTTGAC | 1961 | |
| Qy | 2041 | CACAAAACACCATCAACA | CAGAGGACCAGCCACCTTTGGCTTACCGTCAATTTGNAAGT | 2100 | |
| Db | 1962 | GTCAATGGGAGTT | -----TGTTTTGGCA | CAAAATCAACGGGACTT | 2002 |
| Qy | 2101 | TTTCAGAGTCCACCACTGTCTT | CACGGCCAGTGTCTTATAGACAGGAATGGGAAGGAGG | 2160 | |
| Db | 2003 | TCCAAATGTCGTAAC | ---AATCCGGCCCATTTGACGCAATGGCGGTAGCGGTGACG | 2059 | |
| Qy | 2161 | TCCTGAAGACGATGTGGCTGCTGCGGTCAAAGTGT | TAATGACATTTGTTGATGACTGGAAG | 2220 | |
| Db | 2060 | GTGGGAGGTCTATAT | --AAGCAGAGCTCTCTGGCTAACTAGAGAACCCACTGCTTACTCG | 2117 | |
| Qy | 2221 | CTACCGGTCGGCATCAACATCTTCACTCGCTCGGC | CACAGAGGAGTGAAGTGCAGTG | 2280 | |
| Db | 2118 | CTTATCGAAATTAATACGAC | -----TCACTATAGGAGACCCAAAGCTGGCTAGCGTTTAA | 2173 | |
| Qy | 2281 | ACCAAGGTCTCTCTGGACTCCAGGTG | AAAAAGGAGATAGAGGCCCTCTGCGACAAATGG | 2340 | |
| Db | 2174 | CTTAAAGCTTGATACCGAGCTCGGATCCACTAGT | CCAGTGGTGGAAATCTCGCAGATATC | 2233 | |
| Qy | 2341 | TATACAGGCTTTCCAGGTCTAATAGGTA | CTCCAGGTCTTAAAGTGTATCGGGGAGTCT | 2400 | |
| Db | 2234 | CAGCACAGTGGCGCCGCTCGAT | AAAAATGTCTCAGAGCAACCCGGGAGCTGGTGGTTGAC | 2293 | |
| Qy | 2401 | CTGGTTTACTCTGGAGTTTCGAGGAT | TCAGGACCAATGGGG--AAGACCGGGAAGCCAGG | 2458 | |
| Db | 2294 | TTTTCTCTCCTCAAGCTCTCCGAAAGGATACAGT | GGAGTCAAGTTTACGATGTCGAA | 2353 | |
| Qy | 2459 | ACTTAATGGA | CAAAAAGGCCAGAGGAGAAAAAGGAGTGGAGCATGCAAAAGACAATC | 2518 | |
| Db | 2354 | GAGAA | CAGGACTGAAGCCCCAGAAAGAACTGAAACAGAAAGGAGAGACCCCCAGTGCATC | 2413 | |
| Qy | 2519 | TAATA | CAGTCCGACTGTGTGGTGGCAGCGGCCCTCAGAAAGCAGAGTGGAGATTTTTCA | 2578 | |
| Db | 2414 | AAATGGCAACCATCTCTGGCACTCTGGCGAT | AGCCCGCGGTGAATGGAGCACTGGCCAC | 2473 | |
| Qy | 2579 | CGAAGGCCACTGGGGTACGGTGTGTACGACCGCTGGGA | ACTCCGTGGAGGACTGTGCTGT | 2638 | |
| Db | 2474 | AGCAGCA | GTTTGGATGCGCGGGAGGTACTCCCATGGCAGCACTGAAGCAGCGCTGAGA | 2533 | |
| Qy | 2639 | CTCAGGAGCTTGGGATCAAAAGGTGTTCAAAGTGTGCAT | AAGCAGACTTATTTTGGAAA | 2698 | |
| Db | 2534 | GAGCTGGCGATGAGTTTGAAC | TGCGGTACCGGAGAGCAT----- | 2573 | |
| Qy | 2699 | AGTACGGGTCCAAATATGGCTGAATGAGATATTTTGT | TTTCGGGAAAAGAGTCATCTCATTTGA | 2758 | |
| Db | 2574 | --TCAGTGA | CTAAATCCCAGCTTCATATAACCCCGGAGCAGTATATCAGAGCTTTGA | 2631 | |
| Qy | 2759 | AGAGTCAGCAAAATAGACAGTGGGTGTGAGAGCCTGT | TTCCACAGCAAGAGATCTCTGGGG | 2818 | |
| Db | 2632 | ACAGGTAGTGAATGAATCTTTCCGGATGGGGTAA | ACTGGGGTGCAGTTGTGGCTCTCCTC | 2691 | |
| Qy | 2819 | TCATTTTGCACCTACATAATGCATCATATTTTTCAT | TCACATTTTTTAAACTGTTTATAAAG | 2878 | |
| Db | 2692 | CTCCTTTGGGGGGCACTGTGCGTGGAAAGCGTAGA | CAAGGAGATGCAGGTATTTGTGTAG | 2751 | |
| Qy | 2879 | TGATTTTTTTCCTTTGCTTCACTAAATCAGCTTAAT | TATATTTTAAGAACTTAAGAAAT | 2938 | |
| Db | 2752 | TCGGATTGC | AAAGTTGGATGGCCACTACCTGAAATGACCACTAGAGCTTTGGATCCAGG | 2810 | |
| Qy | 2939 | TTATCCA | CAGAAAAGGAATATTTAAATAACACTTGGATAAAACATATAAAATAGCTTCATAT | 2998 | |
| Db | 2811 | AGAACGGCGCTGGGACACTTTTGTGATCTCTAC | CGGGAACAATACAGCACCCGAGA--- | 2867 | |
| Qy | 2999 | TTGCTTCA | AAATACAGAACCATTTTCAACTCTCTAGGTTTTTAAAGTGGCTCGTGCAGAT | 3058 | |
| Db | 2868 | --GCCG | AAAAGGCAGAGCGTTTCAACCGCT---GGTTCTCAGCGGATGACTGTGGC | 2922 | |

| | | | |
|----|------|--|------|
| QY | 3059 | TGATCCCTCCAGGATATAGTAGTTTCGCTTTTGCCATATAGGAGGGGGAAATGATAGTCTTAT | 3111 |
| Db | 2923 | TGGTGTAGTTCTGCT---GGGCTCACTCTTTCAGTCGGAAAGTGACCAAGATTTCTGCAGATAT | 2979 |
| QY | 3119 | GCAATACTCTTGTAGTCTTGCAACATGGTAAACGATGAGTTAGCAACATGCTTTACAAAGGA | 3178 |
| Db | 2980 | CCATCAACATTCGGGGCGGGAATTCGGGTGAGTCTAGAGGGGCCAGATCTAGATCCGATAC | 3039 |
| QY | 3179 | GAGAAAAGCACCGGTGCATGCCGATTTGGTGGAAAGTAAAGGTGGTATCGATCGTGCTTTATTA | 3238 |
| Db | 3040 | GTCCGACATTGATCCGGCTGTGGAAATGTGTGTCAGTTTAGGGTGTGGAAAGTCCCCAGGCTC | 3099 |
| QY | 3239 | GGAAGCAACAGACGGGTCTGACATGGAATTGGACGAACCACTGAATTCGCGATTCGAGAG | 3298 |
| Db | 3100 | CCCAGCAGGCAGA----- | 3112 |
| QY | 3299 | ATATTGTATTTAAGTGCCTAGCTCGATACGCAAAAGCCCATTTGACCATTTACCAATG | 3358 |
| Db | 3113 | -----AGTATGCAAGCATGC | 3128 |
| QY | 3359 | GTGTGCACCTCCAAGCTTCACGTCGCGCAAGCACCTCAGGGGGCAAGGCGCTGCTAAAGGA | 3418 |
| Db | 3129 | ATCTCAATTAGTCAGCAACCAAGGTGTGGAAAGTCCCAAGCTCCCCAGCAGGCAGGAAGTA | 3188 |
| QY | 3419 | AGCGGAACAGTATAGAAAGCCAGTCCGCAAGAAACGCTGTGACCCCGGATGAATGTCAGCT | 3478 |
| Db | 3189 | TGCAAAGC-----ATGCATCTCAATT | 3209 |
| QY | 3479 | ACTGGGCTATCTGGACAAAGGGAAGAAACGCAAGCGCAAGAGAAGACAGGTAGCTTCGAGTG | 3538 |
| Db | 3210 | AGTCAGCAACCATAGTCCCGGCCCTTAACCTCCGCCCATCCCGCCCTTAACTCCGCCAGTT | 3269 |
| QY | 3539 | GGCTTACATGGCGATAGCTAGACTGGCGGCTTTATGGAACAGCAAGCAACCGGAATTGC | 3598 |
| Db | 3270 | CCGCCCATTTCTCGCCCCATGGCTGACTAATTTTTTTTATTTATGACAGAGGCCGA----- | 3324 |
| QY | 3599 | CAGCTGGGGCGCCCTCTGTGTAAAGTTGGGAAGCCCTGCAAG--TAAACTGGATGCGTTT | 3656 |
| Db | 3325 | -GGCGGCTCGGCCCTCTGAGCTATTCCAGAAGTAGTGAGAGGCTTTTTCGAGGCGCTAG | 3383 |
| QY | 3657 | CTTTCGCCCAAGGATCTGATGGCGCAGGGGATCMAGATCTGATCAAGACACAGGATGAGG | 3716 |
| Db | 3384 | GCTTTTGCAAAAAGCTTGGGCTGCAGGTGCAGGCGGATCTGATCAAGACACAGGATGAGG | 3443 |
| QY | 3717 | ATCGTTTCGATGATGAAACAGATGGATTGCAACGAGGTCTTCGCGCGCTTGGGTGGA | 3776 |
| Db | 3444 | ATCGTTTCGATGATGAAACAGATGGATTGCAACGAGGTCTTCGCGCGCTTGGGTGGA | 3503 |
| QY | 3777 | GAGGCTATTCCGGCTATGACTGGGCACAAACAGACATTCGGCTGCTCTGATCCGCGCGTGT | 3836 |
| Db | 3504 | GAGGCTATTCCGGCTATGACTGGGCAACACAGACATTCGGCTGCTCTGATGCGCGCGTGT | 3563 |
| QY | 3837 | CCGGCTGTACGCGCAGGGGGCCCCGGTTCTTTTGTCAAGACCGACCTGTCCGGTGCCCT | 3896 |
| Db | 3564 | CCGGCTGTACGCGCAGGGGGCCCCGGTTCTTTTGTCAAGACCGACCTGTCCGGTGCCCT | 3623 |
| QY | 3897 | GAATGAACCTCAGACGACGCGGGCTATCGTGGCTGGCCACGACGGGGTTCCTTGG | 3956 |
| Db | 3624 | GAATGAACCTCAGACGACGCGGGCTATCGTGGCTGGCCACGACGGGGTTCCTTGG | 3683 |
| QY | 3957 | CGCAGCTGTCTCGACGTTGTCTACCTGAAGCGGGAAGGGACTGGCTGCTATTGGGGGAAGT | 4016 |
| Db | 3684 | CGCAGCTGTCTCGACGTTGTCTACCTGAAGCGGGAAGGGACTGGCTGCTATTGGGGGAAGT | 3743 |
| QY | 4017 | GCGGGGACGAGATCTCTGTCTCATCTCACTTGTCTCTGCGGAGAAAGTATCCATCATGCG | 4076 |
| Db | 3744 | GCGGGGACGAGATCTCTGTCTCATCTCACTTGTCTCTGCGGAGAAAGTATCCATCATGCG | 3803 |
| QY | 4077 | TGATGCAATGCGGGGCTGCATACGTTGATCCGGTACTCGCCATTTCGACCCCAAGC | 4136 |
| Db | 3804 | TGATGCAATGCGGGGCTGCATACGTTGATCCGGTACTCGCCATTTCGACCCCAAGC | 3863 |
| QY | 4137 | GAACAATTCGATCGAGCGACGACGTACTCGGATGGAAAGCCGGTCTTGTGCTCAGATGGA | 4196 |

| | | | | |
|----|------|--|-----------------------------------|------|
| Db | 3864 | GAACAATCGCATCGAGGAGACAGTA | CTCGAATGAAGCCGGTCTTGTCGATCAGGATGA | 3922 |
| Qy | 4197 | TCTGGAAGAAAGAGCATCAGGGGCTCGCGCCAGAGCCGAAC | CTGTCGCCAGGCTCAAGGCGCG | 4256 |
| Db | 3924 | TCTGGAAGAAAGAGCATCAGGGGCTCGCGCCAGAGCCGAAC | TGTTGCCAGAGCTCAAGGCGCG | 3983 |
| Qy | 4257 | CATGCCCGAGCGGAGGATCTCGTGTGAACCAATGCGGAT | GCCTGCTTGCCGAATATCAT | 4316 |
| Db | 3984 | CATGCCCGAGCGGAGGATCTCGTGTGAACCAATGCGGAT | TGCGATGCGCTTGCCGAATATCAT | 4043 |
| Qy | 4317 | GGTGGAAAATGGCGCTTTCTGGAATCATCGACTGTGGCCGCT | TGGGTGCGCGACCG | 4376 |
| Db | 4044 | GGTGGAAAATGGCGCTTTCTGGAATCATCGACTGTGGCCGCT | TGGGTGCGCGACCG | 4103 |
| Qy | 4377 | CTATCAGGACATAGCGTTGGCTACCCCGTGATATTGCTGAAGAGCT | TGGCGCGCAATGGGC | 4436 |
| Db | 4104 | CTATCAGGACATAGCGTTGGCTACCCCGTGATATTGCTGAAGAGCT | TGGCGCGCAATGGGC | 4163 |
| Qy | 4437 | TGACCGCTTCTCTGTGCTTTACGGATTCGCGCTTCGCGAT | TCGAGCGCATGCCCTTCTA | 4496 |
| Db | 4164 | TGACCGCTTCTCTGTGCTTTACGGATTCGCGCTTCGCGAT | TCGAGCGCATGCCCTTCTA | 4223 |
| Qy | 4497 | TCGCTCTTTGACGAGTTCTTCTGAGCGGACCTCGGGTTCGAT | TAATAAAGATTTT | 4556 |
| Db | 4224 | TCGCTCTTTGACGAGTTCTTCTGAGCGGACCTCGGGTTCGAT | TAATAAAGATTTT | 4283 |
| Qy | 4557 | ATTTAGTCTCCAGAAAAGGGGGAATGAAGAGCCCA | CCTGTAGGTTTGGCAAGCTAGC | 4616 |
| Db | 4284 | ATTTAGTCTCCAGAAAAGGGGGAATGAAGAGCCCA | CCTGTAGGTTTGGCAAGCTAGC | 4343 |
| Qy | 4617 | TTAAGTAAACCCCAATTTTGAAAGGATGGAATAATA | CATAACTGAGAATAGAGAAGTTCA | 4676 |
| Db | 4344 | TTAAGTAAACCCCAATTTTGAAAGGATGGAATAATA | CATAACTGAGAATAGAGAAGTTCA | 4403 |
| Qy | 4677 | ATCAAGTCAAGAACAGATGGAACAGTGAATATGGSCAA | ACAGGATATCTGTGTTAG | 4736 |
| Db | 4404 | ATCAAGTCAAGAACAGATGGAACAGTGAATATGGSCAA | ACAGGATATCTGTGTTAG | 4463 |
| Qy | 4737 | CAGTTCTCGCCCGGCTCAGGGCCAAGAACAGATGGAA | CAGCTGAATATGGGCCAAACAG | 4796 |
| Db | 4464 | CAGTTCTCGCCCGGCTCAGGGCCAAGAACAGATGGAA | CAGCTGAATATGGGCCAAACAG | 4523 |
| Qy | 4797 | GATATCTGTGGTAAGCAGTTCTTGCCCGGCTCAGGGCCA | AGAACAGATGGTCCCAGAT | 4856 |
| Db | 4524 | GATATCTGTGGTAAGCAGTTCTTGCCCGGCTCAGGGCCA | AGAACAGATGGTCCCAGAT | 4583 |
| Qy | 4857 | GCGGTCAGCCCTCAGCAGTTCTTAGAGACCATCAGATG | TTTTCAGGGTGCCTCAGGA | 4916 |
| Db | 4584 | GCGGTCAGCCCTCAGCAGTTCTTAGAGACCATCAGATG | TTTTCAGGGTGCCTCAGGA | 4643 |
| Qy | 4917 | CCTGAAATGACCTGTGCCCTTATTGAACTAAACCAAT | CAGTTGCTTCTCGCTTCTGTTC | 4976 |
| Db | 4644 | CCTGAAATGACCTGTGCCCTTATTGAACTAAACCAAT | CAGTTGCTTCTCGCTTCTGTTC | 4703 |
| Qy | 4977 | GCGCGCTTCTGCTCCCGAGCTCAATAAAGAGCCCA | ACCCCTCACTCGGGGCGCCAG | 5036 |
| Db | 4704 | GCGCGCTTCTGCTCCCGAGCTCAATAAAGAGCCCA | ACCCCTCACTCGGGGCGCCAG | 4763 |
| Qy | 5037 | TCCTCCGATTTGACTGAGTCGCGCGGTACCGGTGAT | CCAAATAAACCCTTTCGAGTTGC | 5096 |
| Db | 4764 | TCCTCCGATTTGACTGAGTCGCGCGGTACCGGTGAT | CCAAATAAACCCTTTCGAGTTGC | 4823 |
| Qy | 5097 | ATCCGACTTGTGGTCTCGCTGTTCTTGGGAGGGT | CTCCTCTGAGTGATTTGACTACCCGT | 5156 |
| Db | 4824 | ATCCGACTTGTGGTCTCGCTGTTCTTGGGAGGGT | CTCCTCTGAGTGATTTGACTACCCGT | 4883 |
| Qy | 5157 | CAGCGGGGCTCTTTCATTTGG | 5177 | |
| Db | 4884 | CAGCGGGGCTCTTTCATTTGG | 4904 | |

RESULT 7

Db 2685 TTCCTGATGTTTTCGACCAAGTGGTGAATTCAGTGGATCCACTAGTAAGCGCCGCA 2744
Qy 2699 AGGTACGGGTCCAAATATGCGTGAATGAAGTATTTTGTTCGGGAAAGAGTCAATCCATGA 2758
Db 2745 GTGTGCTGGAATTAATTCGCTGTCTGCGAGGCGCGGTGTTGGGTGAGTACTCCCTCTC 2804
Qy 2759 AGAGTGCAGAAATTAGACAGTGGGGTGTGAGAGCTGTTCGCACGACGAGAGTGTGGGG 2818
Db 2805 AAA-----AGCGGGCATGACTTTCGCGCTAAGAT-----TG 2835
Qy 2819 TCACCTTTCGACCTACATAATCATCATATTTTCACTTCAATTTTAACTGTTATAAAG 2878
Db 2836 TCAGTTTCCAAAACGAGGAGGATTTGATATCA-----CCTGGCCCGCGGTGATG 2886
Qy 2879 TGATTTTTTTCCTTGTCTTCACTAAAATCAGCTTAATTAATTAATTAAGAACTAAGAAAT 2938
Db 2887 CCTTTGAGGGTGGCGCGTCCATCTGGTTCAGAAAGACAAATCTTTTGTGTCAAGCTTG 2946
Qy 2939 TTATCCACAGAAAGGAATATTTAAATCACTCTGATTAACATATAAATAGCTTCATAT 2998
Db 2947 AGGTGTGCGAGGCTTGAGATCTGGCCATACACTTGAAGTGAC-----AATGACATCCACT 3000
Qy 2999 TTGCTTCAAAATACAGAACCAATTTCAACTTCTTAGGTTTTTAAAGTGTCTGTCGCGAAT 3058
Db 3001 TTGCTTTCTCTCACAGGTGTCACTCCAGGTCCAACTGACAGTTCATCGAGCATGCA 3060
Qy 3059 TGATCCCTCAGGATATAGTAGTTTTCGCTTTTGTATAGGAGGGGGAATAGTCTTTAT 3118
Db 3061 TCTAGGGGGGCAATTCGCCCCCTCTCCCTCCCCCCCCCTAACGTACTGSCCGAAGCG 3120
Qy 3119 GCAATACTTCTGTAGTCTTGCAACATGATACGATAGTGTAGCAACATGCTTCAAGGA 3178
Db 3121 CTTGGAATAAGCGCGGTGTGTGTGTCTATATGTGATTTTCCACCAT----- 3168
Qy 3179 GAGAAAAGCACGTCGATGCGGATTTGTTGGAAGTGAAGTGGTAGCTGTCCTTATTA 3238
Db 3169 -----ATTGCGCTCTTTTGGCAATGTGAGGCGCCGGAACCTTGGCCCTGTCTTCTTG 3220
Qy 3239 GGAAGGCAACAGACGGGTCTGACATGATTTGGAAGAACCACTGAATTCGCAATTCAGAG 3298
Db 3221 ACGAGCATTCCTAGGGGTCTTTCCCTCTCGCCAAAGGAATGCAAGGTCTGTTGAATGTC 3280
Qy 3299 ATATTGTATTAAGTGTCTAGCTCGATACAGAAAGCGCATTTGACCAATTCACCAATG 3358
Db 3281 GTGAAGGAAGCAGTTCTCTCGAAAGCTTCTTGAAGACAAAACAAGCTCTGTAGCACCCTT 3340
Qy 3359 GTGTGCACCTCCAGCTTCACTGCGCAGCACTCAGGCGCAGAGGCTGTAAAGGA 3418
Db 3341 TGCAGGACAGGGAACCCCCACCTGGGACAGGTGCTCTCGCGCCAAAAGCCACGTGTA 3400
Qy 3419 AGCGGAACAAGTACAGTCCGCAAGAACGCTGTGACCCCGGATGAATGTACGCT 3478
Db 3401 TAAGATACACTGCAAGCGGCAC-----AACCCAGTGCCACGTTGTGAGTTGGATAG 3455
Qy 3479 ACTGGGTATCTGGAAGGGAACCGCAAGCGCAAGAGAAAGAGAGTGTGCAAGT 3538
Db 3456 TTGTGGAAGAGTCAATAGGTCTCTCTCAAGCG---TAGTCAACAAGGGGCTGAAGGAT 3511
Qy 3539 GGCTTACATGGGATAGTACACTGGGCGGTTTTATGACAGACAGCAAGCAAGCAAGTTCG 3598
Db 3512 GCCCAGAAAGTACCCCATTTGATGGGAATCTGATCTGGGGCTCTGGTGCACATGCTTTAC 3571
Qy 3599 CAGCTGGGGGCCCTCTGTAAGGTGGGAAGCCCTGCAAGTAACTGATGCTTCT 3658
Db 3572 ATGTGTTTAGTCAGGTTAAAGAGCTTAGGCCCCCGGAA--CCACGGGACGTGGTTT 3629
Qy 3659 TGCGCCAAGGATCTGATGGGCGAGGGGATCAAGATCTGATCAAGAGACAGGATGAGGAT 3718
Db 3630 TCCTTTGAAGAACACGATGATGAAGCTTGCACAAACCCCGGATATTCCTCGACCAATA 3689
Qy 3719 CGTTTCGATGATGAACAAGATGATTTGACGCGAGGTTCTCCGCGCCGCTTGGGTGGAGA 3778
Db 3690 TGGGATCGGCATTTGAAACAAGATGGATTCACGAGGTTCTCCGCGCCGCTTGGGTGGAGA 3749

Qy 3779 GGCTATTTCCGGTATGACTGGGCAACAACAGACAATCGCGTGTCTGTATGCGCCGCTGTTC 3838
Db 3750 GGCTATTTCCGGTATGACTGGGCAACAACAGACAATCGCGTGTCTGTATGCGCCGCTGTTC 3809
Qy 3839 GGCTGTACAGCGAGGCGGCCGCTTCTTTTGTCAAGACCGACTGTCCGCTGCCCTGA 3898
Db 3810 GGCTGTACAGCGAGGCGGCCGCTTCTTTTGTCAAGACCGACTGTCCGCTGCCCTGA 3869
Qy 3899 ATGAACTGCAGGACGAGCGCGCTATCTGTGGCTGGCCACAGCGCGCTTCTTTGCG 3958
Db 3870 ATGAACTGCAGGACGAGCGCGCTATCTGTGGCTGGCCACAGCGCGCTTCTTTGCG 3929
Qy 3959 CAGCTGTCTCGACGTTGTCTACTGAAAGCGGAAAGGACTGCTGCTATTTGGGCGAAGTGC 4018
Db 3930 CAGCTGTCTCGACGTTGTCTACTGAAAGCGGAAAGGACTGCTGCTATTTGGGCGAAGTGC 3989
Qy 4019 CGGGGACAGGATCTCTCTCATCTCACCTGTCTCTGCGGAGAAAGTATCCATCATGCGTG 4078
Db 3990 CGGGGACAGGATCTCTCTCATCTCACCTGTCTCTGCGGAGAAAGTATCCATCATGCGTG 4049
Qy 4079 ATGCAATTCGCGCGCTGCATACGCTTGCATCGGCTACCTGCCATTCGACCAACCAAGCGA 4138
Db 4050 ATGCAATTCGCGCGCTGCATACGCTTGCATCGGCTACCTGCCATTCGACCAACCAAGCGA 4109
Qy 4139 AACATCGCATCGACGAGCACTACTCGGATGGAAGCGGCTTTGTGATCAGATGATC 4198
Db 4110 AACATCGCATCGACGAGCACTACTCGGATGGAAGCGGCTTTGTGATCAGATGATC 4169
Qy 4199 TGAACGAAGACATCAGGCGCTCGCGCAGCGCACTGTTGCGCAGGCTGAAGGCGCA 4258
Db 4170 TGAACGAAGACATCAGGCGCTCGCGCAGCGCACTGTTGCGCAGGCTGAAGGCGCA 4229
Qy 4259 TGCCCGACGCGAGGATCTCTGTGTGACCCATGCGCATGCTGTGCGCAATATCATGG 4318
Db 4230 TGCCCGACGCGAGGATCTCTGTGTGACCCATGCGCATGCTGTGCGCAATATCATGG 4289
Qy 4319 TGAATAATGCGCGCTTTTCTGGAATTCATCGACTGTGCGCGCTGGGTGTGCGCAACGCT 4378
Db 4290 TGAATAATGCGCGCTTTTCTGGAATTCATCGACTGTGCGCGCTGGGTGTGCGCAACGCT 4349
Qy 4379 ATCAGGACATAGCTTGGCTACCGCTGATATTTGCTGAAGACTTGGCGCGCAATGGGCTG 4438
Db 4350 ATCAGGACATAGCTTGGCTACCGCTGATATTTGCTGAAGACTTGGCGCGCAATGGGCTG 4409
Qy 4439 ACCGCTTCTCGTCTTTTACGCTATCGCGCTCCGCTTCGACGCGATCGCTTCTATC 4498
Db 4410 ACCGCTTCTCGTCTTTTACGCTATCGCGCTCCGCTTCGACGCGATCGCTTCTATC 4469
Qy 4499 GCCTTCTTGAAGGTTT-----TTC 4518
Db 4470 GCCTTCTTGAAGGTTTCTGCTCGAGCGGATCTGATCAAGACAGGATGAGGATCGTTT 4529
Qy 4519 TGAGCGGACATCTGGGTTTGAATAAATAAAGATTTTATTTAGTCTCCAGAAAAGGG 4578
Db 4530 CGCGCGGACTCTGGGTTTGAATAAATAAAGATTTTATTTAGTCTCCAGAAAAGGG 4589
Qy 4579 GGAATGAAGACCCCACTGTAGGTTTGGCAAGCTAGCTTAAGTAAAGCCATTTTGAAG 4638
Db 4590 GGAATGAAGACCCCACTGTAGGTTTGGCAAGCTAGCTTAAGTAAAGCCATTTTGAAG 4649
Qy 4639 GCATGGAAAAATAATTAATCAGAAATAGAGAAGTTTCAGATCAAGGTCAGGAACAGATGGA 4698
Db 4650 GCATGGAAAAATAATTAATCAGAAATAGAGAAGTTTCAGATCAAGGTCAGGAACAGATGGA 4709
Qy 4699 ACAGCTGAATATGAGGCGCAACAGAGATATCTGTGTGAAGAGTTCCTGCGCCGCTCAGGG 4758
Db 4710 ACAGCTGAATATGAGGCGCAACAGAGATATCTGTGTGAAGAGTTCCTGCGCCGCTCAGGG 4769
Qy 4759 CCAGAACAGATGGAACAGCTGAATATGGGCAACAGGATATCTGTGTGAAGAGTTCCTGCGCCGCTCAGGG 4818
Db 4770 CCAGAACAGATGGAACAGCTGAATATGGGCAACAGGATATCTGTGTGAAGAGTTCCTGCGCCGCTCAGGG 4829


```
Qy 4819 TGCCCCGGCTCAGGGCCCAAGAACAGATGGTCCCCCAGATGCGGTCCAGCGCTCCAGCCCTCAGCAGTGT 4878
Db 4830 TGCCCCGGCTCAGGGCCCAAGAACAGATGGTCCCCCAGATGCGGTCCAGCGCTCCAGCAGTGT 4899
Qy 4879 CTAGAGAACCAATCAGATGTTTCCAGGGTGCCTCAAGGACCTGAAATGACCTGTGCTTGA 4938
Db 4890 CTAGAGAACCAATCAGATGTTTCCAGGGTGCCTCAAGGACCTGAAATGACCTGTGCTTGA 4949
Qy 4939 TTTGAACCTAACCAATCAGTTCGCTTCTCGCTTCTGTTTCGCGCGCTTCTGCTCCCGGAGCT 4998
Db 4950 TTTGAACCTAACCAATCAGTTCGCTTCTCGCTTCTGTTTCGCGCGCTTCTGCTCCCGGAGCT 5009
Qy 4999 CAATAAAGAGCCCAACACCCCTCACTCGGGCGCCAGTCCCTCCGATGACCTGAGTCCGC 5058
Db 5010 CAATAAAGAGCCCAACACCCCTCACTCGGGCGCCAGTCCCTCCGATGACCTGAGTCCGC 5069
Qy 5059 CGGGTACCCGTGTATCCAAATAAACCTCTTGAGTTCGATCCGACTTGTGTGCTCGCTGT 5118
Db 5070 CGGGTACCCGTGTATCCAAATAAACCTCTTGAGTTCGATCCGACTTGTGTGCTCGCTGT 5129
Qy 5119 TCCCTGGAGGGTCTCCTCTGAGTGAATGACTACCCGTCAGCGGGGGTCTTTCATTTGG 5177
Db 5130 TCCCTGGAGGGTCTCCTCTGAGTGAATGACTACCCGTCAGCGGGGGTCTTTCATTTGG 5188

RESULT 8
BD138564 7165 bp DNA linear PAT 18-SEP-2002
LOCUS Expression system of nucleotide with reduced immunogenicity for use
DEFINITION in gene therapy.
ACCESSION BD138564
VERSION BD138564.1 GI:23233509
KEYWORDS JP 2002508976-A/3.
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 7165)
AUTHORS Radosevich, T.J. and Jr, C.J.L.
TITLE Expression system of nucleotide with reduced immunogenicity for use
JOURNAL Patent: JP 2002508976-A 3 26-MAR-2002;
COMMENT HUMAN GENE THERAPY RESEARCH INSTITUTE
OS Artificial Sequence
PN JP 2002508976-A/3
PD 26-MAR-2002
PF 13-JAN-1999 JP 2000540263
PR 14-JAN-1998 US 60/071409
PI THOMAS J RADOSEVICH, CHARLES J LINK JR
PC C12N15/09,A61K31/711,A61K48/00,A61P37/06,C12N1/15,C12N1/19, PC
C12N1/21,
PC C12N5/10/A61K35/76, (C12N15/09,C12N1/93),C12N15/00,C12N5/00,
PC (C12N15/00,C12N1/93)
CC Description of Artificial Sequence : recombinant vector FH
Key Location/Qualifiers
FT source 1. .7165
FT /organism='Artificial Sequence'.
FEATURES
source Location/Qualifiers
1. .7165
/organism='synthetic construct'
/mol_type='genomic DNA'
/db_xref='taxon:32630'
ORIGIN
Query Match 31.8%; Score 1644.6; DB 6; Length 7165;
Best Local Similarity 62.9%; Pred. No. 0;
Matches 3284; Conservative 0; Mismatches 1309; Indels 624; Gaps 23;
Qy 1 TTTGAAGACCCCGTAGGTGCGAGCTAGCTTAAGTAACGCCACTTTGCAAGGCAT 60
Db 175 TTTGAAGACCCCGTAGGTGCGAGCTAGCTTAAGTAACGCCACTTTGCAAGGCAT 234
Qy 61 GGAATAATACATAACTGAGAAATAGAAAAGTTTCAGATCAAGGTCAGGAAACAAAGAAACAGC 120
```

```
Db 235 GGAATAATACATAACTGAGAAATAGAAAAGTTTCAGATCAAGGTCAGGAAACAAAGAAACAGC 294
Qy 121 TGAATACCAACACAGGATATCTGTGTAAGCGGTTCTTGCCTCCCGCTCAGGCCCAAGACACA 180
Db 295 TGAATACCAACACAGGATATCTGTGTAAGCGGTTCTTGCCTCCCGCTCAGGCCCAAGACACA 354
Qy 181 GATGAGACAGCTGAGTGTATGGGGCCAAACAGGATATCTGTGTAAGCGGTTCTTGCCTCCCG 240
Db 355 GATGAGACAGCTGAGTGTATGGGGCCAAACAGGATATCTGTGTAAGCGGTTCTTGCCTCCCG 414
Qy 241 CTGGGGCCCAAGAACAGATGTTCCCAAGATGCGGTCCAGCCCTCAGCAGTTCCTAGTGA 300
Db 415 CTGGGGCCCAAGAACAGATGTTCCCAAGATGCGGTCCAGCCCTCAGCAGTTCCTAGTGA 474
Qy 301 TCATCAGATGTTTCCAGGGTGCCTCAAGGACCTGAAAATGACCTGACCTTATTTGAAAC 360
Db 475 TCATCAGATGTTTCCAGGGTGCCTCAAGGACCTGAAAATGACCTGACCTTATTTGAAAC 534
Qy 361 TAACCAATCAGTTCGCTTCTCGCTTCTGTTCGCGCGCTTCCGCTCTCCGAGCTCAATAAA 420
Db 535 TAACCAATCAGTTCGCTTCTCGCTTCTGTTCGCGCGCTTCCGCTCTCCGAGCTCAATAAA 594
Qy 421 AGAGCCCAACACCCCTCACTCGGGCGCCAGTCTTCCGATAGACTGCGTCCCGGGTAC 480
Db 595 AGAGCCCAACACCCCTCACTCGGGCGCCAGTCTTCCGATAGACTGCGTCCCGGGTAC 654
Qy 481 CCGTATTTCCCAATAAAGCCCTTGTGTTTGCATCCGAATCGTGGTCTCGCTGTTCCCTTG 540
Db 655 CCGTATTTCCCAATAAAGCCCTTGTGTTTGCATCCGAATCGTGGTCTCGCTGTTCCCTTG 714
Qy 541 GAGGGGTCTCTCTCTGAGTGAATGACTACCAACGAGGGGGTCTTTCATTTGGGGGCTCGT 600
Db 715 GAGGGGTCTCTCTCTGAGTGAATGACTACCAACGAGGGGGTCTTTCATTTGGGGGCTCGT 774
Qy 601 CCGGGATTTGAGACCCCTCCCGAGGGACCAACGACCCACACCGGAGGTAACTGGCC 660
Db 775 CCGGGATTTGAGACCCCTCCCGAGGGACCAACGACCCACACCGGAGGTAACTGGCC 834
Qy 661 AGCAACTATCTGTGCTGTCCGATGCTAGTGTCTATGTTGATCTATGCGCCCTGCG 720
Db 835 AGCAACTATCTGTGCTGTCCGATGCTAGTGTCTATGTTGATCTATGCGCCCTGCG 894
Qy 721 TCTGTACTAGTTAGCTAACTAGCTCTGTATCTGCGGACCCGCTGGTGGAACTGACGAGTT 780
Db 895 TCTGTACTAGTTAGCTAACTAGCTCTGTATCTGCGGACCCGCTGGTGGAACTGACGAGTT 954
Qy 781 CTGAACACCCCGCGCAACCTCGGAGACGTCCTCAGGAGCTTTGGGGGCGCTTTTGTGG 840
Db 955 CTGAACACCCCGCGCAACCTCGGAGACGTCCTCAGGAGCTTTGGGGGCGCTTTTGTGG 1014
Qy 841 CCGGACCTGAGGAGGGAGTGGATGGAATCCGACCCCGCTCAGGATATGTTGTTCTGTT 900
Db 1015 CCGGACCTGAGGAGGGAGTGGATGGAATCCGACCCCGCTCAGGATATGTTGTTCTGTT 1074
Qy 901 AGGAGACGAGAACTTAAACAGTTCCCGCTCTGAAATTTTTCGTTTCGGTTTGGAA 960
Db 1075 AGGAGACGAGAACTTAAACAGTTCCCGCTCTGAAATTTTTCGTTTCGGTTTGGAA 1134
Qy 961 CCGAAGCCCGCGCTTGTCTGTCTGACGACCAAGCTTGGGCTGCGAGTCTAGAGGA 1020
Db 1135 CCGAAGCCCGCGCTTGTCTGTCTGCGCAG----- 1162
Qy 1021 TCAATTCGGCACGAGTAAATCGGTGCTGCGCTTTTAGGACATATAGATATGSCACAGT 1080
Db 1163 -----CGCTGACGATCGTTCTGTGTGTCTGTCTGTCTGCTGAGTGTGT 1202
Qy 1081 GGGATGACTTTCTGTATCAGCAAGAGGACACTGACAGCTGTACAGAGTCTGTGAAGTTCG 1140
Db 1203 TTCTGTATTTGTCTGAAAATTAGGCCAGACTGTATCC-----ACTCCCTTAAGTTG 1255
Qy 1141 ATGCTCGCTCAGTGACAGCTTTGCTTCTCCCATCTTAAATAATGGGCCCACTCTTCAAG 1200
Db 1256 ACCTTAGGTCACTGGAAAGATGTGAGCGGATCGCTCAACCAACGAGTCCGTAGATGTCAAG 1315
```


Db 3003 GCCAAAGCCAGGTGTATAGATACACCTGCAAGGGCGGCACA----- 3045
Qy 3421 CGGAACAGTGAAGCCAGTCCGAGAAACGGTGTGACCCCGGATGAATGTGAGTAC 3480
Db 3046 ---ACCCAGTGCACAGTGTGTGATAGTTGTGGAAGAGTCAATGGCTCTCCTC 3102
Qy 3481 TGGGCTATCTGGACAAGGGAAACCGCAAGCCGCAAGAGAGAGAGTGTGAGTGGG 3540
Db 3103 AAGGTAGTCAACAGGGGCTGAAGAGTCCAGAGAGTACCCATTGTATGGGAATCTG 3162
Qy 3541 CTACATGGCGATAGCTAGACTGGCGGTTTATGGACAGCAAGCGAACCCGGAATTGCCA 3600
Db 3163 AT--CTGGGCTCGGTGCACATGCTTTACATGTGTTTAGTTCGAGGTTAAAAAGCTTA 3220
Qy 3601 GCTGGGGGCGCTCTGTGAAGTGGGAGCCCTGCAAAAGTAAACTGATGGCTTTCTTG 3660
Db 3221 GGCCCGCCGAACCAAGGGGAGTGGTTTCTTTGAAAAACACGATGAAGCTTGCCAC 3280
Qy 3661 CCGCAAGGATCTGATGGCGCAGGGATCAAGATCTGATCAAGAGACAGGATGAGGATCG 3720
Db 3281 AACCCGGGAT-----AATCTGCGCAATATGGGATCGG----- 3317
Qy 3721 TTTCGATGATTGAACAGATGATGTCACGCGAGTTCCTCGGCGCTTGGGTGGAGAG 3780
Db 3318 -----CCATTGAACAAGATGGATGCAAGCAGGTTCTCGGCGGCTTGGGTGGAGAG 3370
Qy 3781 CTATTGGCTATGATGGGCAACAAGACAATCGGCTGCTGTGATGCGCGCTGTTCCGG 3840
Db 3371 CTATTGGCTATGATGGGCAACAAGACAATCGGCTGCTGTGATGCGCGCTGTTCCGG 3430
Qy 3841 CTGTGCGCAGGGGCGCGGTTCTTTTGTCAAGACCGACCTGCGGTGCCCTGAAT 3900
Db 3431 CTGTGCGCAGGGGCGCGGTTCTTTTGTCAAGACCGACCTGTCGGTGCCTGAAT 3490
Qy 3901 GAACTGACGACGAGGCGCGCTATCGTGGCTGCGCACGCGGCGTTCCTTGGCGCA 3960
Db 3491 GAACTGACGACGAGGCGCGCTATCGTGGCTGCGCACGCGGCGTTCCTTGGCGCA 3550
Qy 3961 GCTGTGTCGAGTGTCTGAAGCGGGAAGGACTGGGCTGTTATTTGGGCGAAGTGGCG 4020
Db 3551 GCTGTGTCGAGTGTCTGAAGCGGGAAGGAGTGGGCTGTTATTTGGGCGAAGTGGCG 3610
Qy 4021 GGGCAGGATCTCTGTGATCTCACCTGCTCCTCGCAGAAAGTATCCATCATGCTGAT 4080
Db 3611 GGGCAGGATCTCTGTGATCTCACCTGCTCCTCGCAGAAAGTATCCATCATGCTGAT 3670
Qy 4081 GCAATGGCGGCTGCTATGATCGGCTACCTGCGCATTCGACCCATTCGACCAAGCGAAA 4140
Db 3671 GCNATGGCGGCTGCTATGATCGGCTACCTGCGCATTCGACCCATTCGACCAAGCGAAA 3730
Qy 4141 CATGCGATTCGAGGAGCAGTACTCGGATGAAAGCGGCTCTGTGATCAGGATGATCTG 4200
Db 3731 CATGCGATTCGAGGAGCAGTACTCGGATGAAAGCGGCTCTGTGATCAGGATGATCTG 3790
Qy 4201 GACGAAGCATCAGGGGCTCGCCGACCGCAACTGTTGCGCAGGCTCAAGGGCGGATG 4260
Db 3791 GACGAAGCATCAGGGGCTCGCCGACCGCAACTGTTGCGCAGGCTCAAGGGCGGATG 3850
Qy 4261 CCCGACGCGCAGGATCTGCTGCTGACCATGGGATGCTGCTTGGCGGAATCATGGTG 4320
Db 3851 CCCGACGCGCAGGATCTGCTGCTGACCATGGGATGCTGCTTGGCGGAATCATGGTG 3910
Qy 4321 GAAAAATGGCCCTTTTCTGGATTCATGATGCTGGCGGCTGGGTGTGGCGGACCGCTAT 4380
Db 3911 GAAAAATGGCCCTTTTCTGGATTCATGATGCTGGCGGCTGGGTGTGGCGGACCGCTAT 3970
Qy 4381 CAGGACATAGGCTTGGCTACCGGTGATATTCGTAAGAGCTTGGCGGGAATGGCTGAC 4440
Db 3971 CAGGACATAGGCTTGGCTACCGGTGATATTCGTAAGAGCTTGGCGGGAATGGCTGAC 4030
Qy 4441 CGCTTCTCTGCTTTTACGGTATCGCGCTATCGCGCTATCGCGATTCGAGCGCATCGCTTCTATCGC 4500

Db 4031 CGCTTCTCTGCTTTACGGTATCGCGCTCCGATTCGAGCGCATCGCTTCTATCGC 4090
Qy 4501 CTTCTTGACGAGTTC-----TTCTG 4520
Db 4091 CTTCTTGACGAGTTCGTGTCGAGCGGATCTGATCAAGAGACAGGATGAGATCGTTTCG 4150
Qy 4521 AGCGGAGCTCTGGGGTTCGATAAAATAAAGATTTTATTAGTCTCCAGAAAAAGGGGG 4580
Db 4151 CGCGGAGCTCTGGGGTTCGATAAATAAAGATTTTATTAGTCTCCAGAAAAAGGGGG 4210
Qy 4581 AATGAAAGACCCACCTGTAGTTGGCAAGCTAGCTTAAGTAACGCAATTTTCAAGGC 4640
Db 4211 AATGAAAGACCCACCTGTAGTTGGCAAGCTAGCTTAAGTAACGCAATTTTCAAGGC 4270
Qy 4641 ATGAAAAATACATTAAGTGAAGATAGAGATTCAGATCAAGGTACAGGACAGATGGAAC 4700
Db 4271 ATGAAAAATACATTAAGTGAAGATAGAGATTCAGATCAAGGTACAGGACAGATGGAAC 4330
Qy 4701 AGCTGAATATGGGCCAAACAGGATATCTGTGTAAGCAGTTCTCTGCCCGGCTCAGGGCC 4760
Db 4331 AGCTGAATATGGGCCAAACAGGATATCTGTGTAAGCAGTTCTCTGCCCGGCTCAGGGCC 4390
Qy 4761 AAGAACAAGTGAACAGCTGAATATATGGGCCAAACAGGATATCTGTGTAAGCAGTTCTCT 4820
Db 4391 AAGAACAAGTGAACAGCTGAATATATGGGCCAAACAGGATATCTGTGTAAGCAGTTCTCT 4450
Qy 4821 CCCGGCTCAGGGCCAAGACAGATGCTGCCAGATGCGGTCCAGCCCTCAGCAGTTTCT 4880
Db 4451 CCCGGCTCAGGGCCAAGACAGATGCTGCCAGATGCGGTCCAGCCCTCAGCAGTTTCT 4510
Qy 4881 AGAACAACATCAGATGTTTCCAGGGTCCCACAGGACCTGAAATGACCTGTGCTTATT 4940
Db 4511 AGAACAACATCAGATGTTTCCAGGGTCCCACAGGACCTGAAATGACCTGTGCTTATT 4570
Qy 4941 TGAACTAACCAATCAGTTGCTTCTGCTTCTGTTCCGCGCTTCTGCTCCCCAGCTCA 5000
Db 4571 TGAACTAACCAATCAGTTGCTTCTGCTTCTGTTCCGCGCTTCTGCTCCCCAGCTCA 4630
Qy 5001 ATAAAGAGCCCAACACCCCTCCTACCTCGGGGCGGCGAGTCTCCGATTCAGTTCGCGCG 5060
Db 4631 ATAAAGAGCCCAACACCCCTCCTACCTCGGGGCGGCGAGTCTCCGATTCAGTTCGCGCG 4690
Qy 5061 GGTACCCGTGTATCCAAATAAACCTCTTGCAGTTGCAATCCGACTTGTGCTCTGCTGTTTC 5120
Db 4691 GGTACCCGTGTATCCAAATAAACCTCTTGCAGTTGCAATCCGACTTGTGCTCTGCTGTTTC 4750
Qy 5121 CTTGGAGGCTCTCTCTGATGATGACTACCGGTACGCGGGGCTTTTCATTGG 5177
Db 4751 CTTGGAGGCTCTCTCTGATGATGACTACCGGTACGCGGGGCTTTTCATTGG 4807

RESULT 9

AR214682 LOCUS 5594 bp DNA linear PAT 25-SEP-2002
DEFINITION Sequence 29 from patent US 6410220.
ACCESSION AR214682
VERSION AR214682.1 GI:23312592
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 5594)
AUTHORS Hodgson,C.P., Zink,M.A. and Xu,G.
TITLE Self-assembling genes, vectors and uses thereof
JOURNAL Patent: US 6410220-A 29 25-JUN-2002;
Nature Technology Corp; Omaha, NE
FEATURES
source 1. 5594
Location/Qualifiers
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 31.5%; Score 1629.4; DB 6; Length 5594;

QY 2281 ACCAAGTCTCTCGGACTCCAGTGTGAAGGAGATAGAGGCCCTCCTGGACAAAATGG 2340
DB |||||
QY 2006 AACAAATATCTGAACCTTTTACCCCGTTGCCGGCA----- 2039
DB |||||
QY 2341 TATACCAGGCTTTCCAGTCTTAATAGTACTCCAGGCTTTAAAGGTGATCGGGGGATCT 2400
DB ----- 2039
QY 2401 CTGGTTTACCTGGAGTTTCGAGGATTTCCAGGACCAATGGGGAGACCGGGAAGCCAGGAC 2460
DB ----- 2039
QY 2461 TTAATGACAAAAGGCCAGAGAGGAGAAAAGGGAGTGGNAGCATGCAAAAGACAACTTA 2520
DB |||||
QY 2040 -----ACGGTCAG 2047
DB -----
QY 2521 ATACAGTCCGACTGGTGGGTGGCAGCGCCCTCACGAAGCAGAGTGGAGATTTTTCACG 2580
DB |||||
QY 2048 GTCTCTCCAAAGTGTCTGACGCAACCCCACTGGATGGGGCTTGGCTATCGGCCATA 2107
DB |||||
QY 2581 AAGGCCAGTGGGGTACGGTGTGTGACGACCGCTGGGAACTGCCGTGGAGGACTGGTCTCT 2640
DB |||||
QY 2108 GCCGATGCGCGGACCTTTGTGGCTCTCTGCCG----- 2141
DB -----
QY 2641 GCAGGAGCTTGGGATACAAAGGTGTTCAAAGTGTGCATTAAGCGAGCTTATTTTGGAAAAG 2700
DB |||||
QY 2142 -----ATCCATATCGGGAACCTCTAGCAGCTGT----- 2171
DB -----
QY 2701 GTACGGGTCCAAATATGCGTGAATGAAGTATTTTGTTCGGGAAGAGTCACTCAATGAAG 2760
DB ----- 2171
QY 2761 AGTGCAAAATTAGACAGTGGGGTGTGAGAGCTGTTCCGACGAGAAAGATGCTGGGGTC 2820
DB |||||
QY 2172 -----TTTGCTGCGACGGCGTCTGGAGCGAA 2197
DB -----
QY 2821 ACTTTGACCTACATAATGATCATATATTTTCAATTCACATTTTTPAAACTGTTATAAGTG 2880
DB |||||
QY 2198 ACTTATCGGC----- 2207
DB -----
QY 2881 ATTTTTTCTTTGCTTCACTAAATCAGCTTAATTAATTAATTAAGAAACTAAGATTTT 2940
DB ----- 2207
QY 2941 ATCCACAGAAAAGGAATATTTAAATAATCACTGGATAAACAATAATAAATAGCTTCATATTT 3000
DB |||||
QY 2208 -----ACCGAACCTCT 2219
DB -----
QY 3001 GCTTCAAATACCAGAACATTTCAACTTCTCTAGGTTTTTAAGTGGCTCGTGCCGAATTG 3060
DB |||||
QY 2220 GTTGTCTCTCTCGGAAATACACCTCTTTTCCATGGCTGCTAGGGTGTGCTGCCAATGG 2279
DB |||||
QY 3061 ATCCCTCAGGATATAGTAGTTGCTTTGCTATGATAGGAGGGGGAATAGTAGTCTTATGC 3120
DB |||||
QY 2280 ATCCCTCAGGATATAGTAGTTGCTTTGCTATGATAGGAGGGGGAATAGTAGTCTTATGC 2339
DB |||||
QY 3121 AATACTCTCTGTAGTCTTGAACATGTTAAAGTGTGGAAGTGTAGCAACATGCTTTTACAGGAGA 3180
DB |||||
QY 2340 AATACACTGTGTGCTTTGCAACATGTTAAAGTGTGGAAGTGTAGCAACATGCTTTTACAGGAGA 2399
DB |||||
QY 3181 GAAAAAGCAACCGTGATCGGATTTGGTGAAGTGTAGGATGATCGTGCCTTTATTAGG 3240
DB |||||
QY 2400 GAAAAAGCAACCGTGATCGGATTTGGTGAAGTGTAGGATGATCGTGCCTTTATTAGG 2459
DB |||||
QY 3241 AAGCAACAGACGGGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3300
DB |||||
QY 2460 AAGCAACAGACAGGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2519
DB |||||
QY 3301 -ATTGATTTTAAAGTGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3357
DB |||||
QY 2520 AATTGATTTTAAAGTGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2579
DB |||||

QY 3358 GGTGTGCACCT--CCAAGCTTACGCTGCCGCAAGCACTCAGGGCGCAAGGGCTGCTAAA 3415
DB |||||
QY 2580 GGTGTGCACCTTCCAAAGCTTCACTGCTGCCGCAAGCACTCAGGGCGCAAGGGCTGCTAAA 2639
DB |||||
QY 3416 GGAAGCGGAACACGTAGAAAGCCAGTCCGAGAAAAGGTGCTGACCCCGGATGAATGTCA 3475
DB |||||
QY 2640 GGAAGCGGAACACGTAGAAAGCCAGTCCGAGAAAAGGTGCTGACCCCGGATGAATGTCA 2699
DB |||||
QY 3476 GCTACTGGGCTATCTGGCAAGGGGAAACGCAAGCGCAAGAGAGAAAGCAGGTAGCTTGA 3535
DB |||||
QY 2700 GCTACTGGGCTATCTGGCAAGGGGAAACGCAAGCGCAAGAGAGAAAGCAGGTAGCTTGA 2759
DB |||||
QY 3536 GTGGGCTTACATCGCATAGTACAGTGGCGGGTTTTTATGACAGAGCAAGCAAGCGAAT 3595
DB |||||
QY 2760 GTGGGCTTACATCGCATAGTACAGTGGCGGGTTTTTATGACAGCAAGCAAGCGAAT 2819
DB |||||
QY 3596 TGCAGCTGGGGCGCCCTCTGGTAAAGTTGGGAAGCCCTGCAAAAGTAAACTGATGCTT 3655
DB |||||
QY 2820 TGCAGCTGGGGCGCCCTCTGGTAAAGTTGGGAAGCCCTGCAAAAGTAAACTGATGCTT 2879
DB |||||
QY 3656 TCTTGGCGCCCAAGGATCTGATGGCGCAGGGGATCAAGATCTGATCAAGAGACAGGATGAG 3715
DB |||||
QY 2880 TCTTGGCGCCCAAGGATCTGATGGCGCAGGGGATCAAGATCTGATCAAGAGACAGGATGAG 2939
DB |||||
QY 3716 GATCGTTTCCCATGATGAAACAAGATGGATGGACGAGGTTCTCCGGCGCTTGGGTGG 3775
DB |||||
QY 2940 GATCGTTTCCCATGATGAAACAAGATGGATGGACGAGGTTCTCCGGCGCTTGGGTGG 2999
DB |||||
QY 3776 AGAGGCTATTTCCGCTATGACTGGGCAACAAGCAATCGGCTGCTGATGATGATGATGATGAT 3835
DB |||||
QY 3000 AGAGGCTATTTCCGCTATGACTGGGCAACAAGCAATCGGCTGCTGATGATGATGATGATGAT 3059
DB |||||
QY 3836 TCCGGCTGTGACGCGCAGGGCGCCCGGTTCTTTTGTGCAAGACCGCATCTGCTCCGGTGCCC 3895
DB |||||
QY 3060 TCCGGCTGTGACGCGCAGGGCGCCCGGTTCTTTTGTGCAAGACCGCATCTGCTCCGGTGCCC 3119
DB |||||
QY 3896 TGAATGAACTGCAAGACGAGGCGCGGCTATGCTGGCTGGGCAACAAGCGGGGTTCCCTT 3955
DB |||||
QY 3120 TGAATGAACTGCAAGACGAGGCGCGGCTATGCTGGCTGGGCAACAAGCGGGGTTCCCTT 3179
DB |||||
QY 3956 GCGCAGCTGTGCTGCAAGTGTGCTGCAAGCGGGAAGGAGTGGCTGCTTATGCGGCGAAG 4015
DB |||||
QY 3180 GCGCAGCTGTGCTGCAAGTGTGCTGCAAGCGGGAAGGAGTGGCTGCTTATGCGGCGAAG 3239
DB |||||
QY 4016 TGCCGGGCGAGGATCTCTGCTCATCTCACCTGCTCTGCCGAGAAAGTATCCATCATGG 4075
DB |||||
QY 3240 TGCCGGGCGAGGATCTCTGCTCATCTCACCTGCTCTGCCGAGAAAGTATCCATCATGG 3299
DB |||||
QY 4076 CTGATGCAATGCGCGGCTGCATACGCTTGAATCCGGCTACCTGCCCATTCGACCCCAAG 4135
DB |||||
QY 3300 CTGATGCAATGCGCGGCTGCATACGCTTGAATCCGGCTACCTGCCCATTCGACCCCAAG 3359
DB |||||
QY 4136 CGAAACATGCGCATGCGAGGACGCTACTCGGATGGAAGCCGCTTGTGCTGATCAGGATG 4195
DB |||||
QY 3360 CGAAACATGCGCATGCGAGGACGCTACTCGGATGGAAGCCGCTTGTGCTGATCAGGATG 3419
DB |||||
QY 4196 ATCTGACCAAGAGCATCAGGGGCTCGCGCAGCGCAACTGTTTCCGAGGCTCAAGCGCG 4255
DB |||||
QY 3420 ATCTGACCAAGAGCATCAGGGGCTCGCGCAGCGCAACTGTTTCCGAGGCTCAAGCGCG 3479
DB |||||
QY 4256 GCATGCCCGCAGCGAGGATCTCGTGTGATGCCATGCGATGCTGCTTCCGGAATATCA 4315
DB |||||
QY 3480 GCATGCCCGCAGCGAGGATCTCGTGTGATGCCATGCGATGCTGCTTCCGGAATATCA 3539
DB |||||
QY 4316 TGGTGAATAATGCGCGCTTTTCTGGATTCATTCGACTGTGGCCGCTGGGTGTGGCGGACC 4375
DB |||||
QY 3540 TGGTGAATAATGCGCGCTTTTCTGGATTCATTCGACTGTGGCCGCTGGGTGTGGCGGACC 3599
DB |||||
QY 4376 GCTATCAGGACATAGGCTTGGCTTACCGCTGATATTCCTGAAGAGCTTGGCGGCAATGGG 4435
DB |||||
QY 3600 GCTATCAGGACATAGGCTTGGCTTACCGCTGATATTCCTGAAGAGCTTGGCGGCAATGGG 3659
DB |||||
QY 4436 CTGACCGCTTCTCTGCTGCTTTTACGGTATCGCCGCTCCCGATTCGACGCGCATGCGCTTCT 4495
DB |||||

|||||
Db 3660 CTGACCGCTTCTCGTGTCTTACGGTATCGCGCTCCGGATTGCGACGCGATCGCCTTCT 3719
Qy 4496 ATCGCCTTCTGACGAGTCTTCTGACGGGACTCTGGGTTTGATA 4542
Db 3720 ATCGCCTTCTGACGAGTCTTCTGACGGGACTCTGGGTTTGAAA 3766
RESULT 11
AX359935
LOCUS AX359935 5130 bp DNA linear PAT 13-FEB-2002
DEFINITION Sequence 9 from Patent WO0202783.
ACCESSION AX359935
VERSION AX359935.1 GI:18675574
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Bleck, G.T.
TITLE Expression vectors
JOURNAL Patent: WO 0202783-A 9 10-JAN-2002;
Gala Design, Inc. (US)
FEATURES
source 1. 5130
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Synthetic"
ORIGIN
Query Match 30.78; Score 1588.6; DB 6; Length 5130;
Best Local Similarity 66.18; Pred. No. 0;
Matches 3007; Conservative 0; Mismatches 754; Indels 786; Gaps 18;
Qy 1 TTTGAAAGACCCCAACCGTAGTGGCAAGCTAGCTTAAGTAAACGCCACTTTTGAAGGCAT 60
Db 1 TTTGAAAGACCCCAACCGTAGTGGCAAGCTAGCTTAAGTAAACGCCACTTTTGAAGGCAT 60
Qy 61 GGAATAATACATACTGAGATAGAAAAGTTTCAGATCAAGGTCAAGGTAAGGAAACAGC 120
Db 61 GGAATAATACATACTGAGATAGAAAAGTTTCAGATCAAGGTCAAGGTAAGGAAACAGC 120
Qy 121 TGATACCAACAGGATATCTGTGTAAGCGTTCTGCGCGCTCAGCGCCCAAGAAC 180
Db 121 TGAATACCAACAGGATATCTGTGTAAGCGTTCTGCGCGCTCAGCGCCCAAGAAC 180
Qy 181 GATGAGACAGCTGAGTGATGGGCCAAACAGGATATCTGTGTAAGCAGTTCCTGCCCCG 240
Db 181 GATGAGACAGCTGAGTGATGGGCCAAACAGGATATCTGTGTAAGCAGTTCCTGCCCCG 240
Qy 241 CTGCGGGCCAAAGAACAGATGTTCCCAAGATGCGGTCCAGCCCTCAGCAGTTTCTAGTGAA 300
Db 241 CTGCGGGCCAAAGAACAGATGTTCCCAAGATGCGGTCCAGCCCTCAGCAGTTTCTAGTGAA 300
Qy 301 TCATCAGATGTTCCAGGGTGGCCCAAGGACCTGAAATGACCTGACCTTATTTGAAC 360
Db 301 TCATCAGATGTTCCAGGGTGGCCCAAGGACCTGAAATGACCTGACCTTATTTGAAC 360
Qy 361 TAAACCAATCAGTTCGCTTCTGCTTCTGTCGCGCTCTCGCTCTCCGAGCTCAATATA 420
Db 361 TAAACCAATCAGTTCGCTTCTGCTTCTGTCGCGCTCTCGCTCTCCGAGCTCAATATA 420
Qy 421 AGAGCCCAACCCCTCACTCGGCGCCAGTCTTCGATAGACTGCGTCCGCGGTAC 480
Db 421 AGAGCCCAACCCCTCACTCGGCGCCAGTCTTCGATAGACTGCGTCCGCGGTAC 480
Qy 481 CCGTATTTCCCAATAAAGCCTTCTGCTTTGTCATCCGAATCGTGGTCTCGCTGTTCTTG 540
Db 481 CCGTATTTCCCAATAAAGCCTTCTGCTTTGTCATCCGAATCGTGGTCTCGCTGTTCTTG 540
Qy 541 GGAGGGTCTCTCTGAGTGATTGACTACCAACGACGGGGTCTTTTCATTTGGGGGCTCGT 600
|||||

Db 541 GGAGGGTCTCTCTGAGTGATTGACTACCAACGACGGGGTCTTTTCATTTGGGGGCTCGT 600
Qy 601 CCGGGATTTGGAGACCCCTGCCAGAGGACCAACGACCCACCGGGAGGTAGAGCTGCGC 660
Db 601 CCGGGATTTGGAGACCCCTGCCAGAGGACCAACGACCCACCGGGAGGTAGAGCTGCGC 660
Qy 661 AGCAACTTATCTGTCTGTCTGTCCGATTTCTAGTGTCTATGTTTATGAGTATGCGCCTGCG 720
Db 661 AGCAACTTATCTGTCTGTCTGTCCGATTTCTAGTGTCTATGTTTATGAGTATGCGCCTGCG 720
Qy 721 TCTGTACTAGTTAGCTAACTAGCTCTGTATCTGGCGACCCGCTGGTGAACCTGACGAGTT 780
Db 721 TCTGTACTAGTTAGCTAACTAGCTCTGTATCTGGCGACCCGCTGGTGAACCTGACGAGTT 780
Qy 781 CTGAACACCGCGCCGCAACCTCTGGAGAGCTGCCAGGACCTTTGGGGGCGCTTTTCTGG 840
Db 781 CTGAACACCGCGCCGCAACCTCTGGAGAGCTGCCAGGACCTTTGGGGGCGCTTTTCTGG 840
Qy 841 CCGGACCTGAGGAAGGAGTTCGATGTGGAATCCGACCCCGCTCAGGATATGTGTTCTGTT 900
Db 841 CCGGACCTGAGGAAGGAGTTCGATGTGGAATCCGACCCCGCTCAGGATATGTGTTCTGTT 900
Qy 901 AGGAGACGAGAACCTTAAACAGTTCCGCTCCGCTCTGAATTTTTCGTTTCGTTGGA 960
Db 901 AGGAGACGAGAACCTTAAACAGTTCCGCTCCGCTCTGAATTTTTCGTTTCGTTGGA 960
Qy 961 CCGAAGCCGCGCTCTTGTCTGCTGAGCAACGCTTGGGCTGAGGCTCAGACTCTAGAGGA 1020
Db 961 CCGAAGCCGCGCTCTTGTCTGCTGAGCAACGCTTGGGCTGAGGCTCAGGCTCAGGGA 1020
Qy 1021 TCAATTCCGCAACGAGTAAATCGGTGCTCTTAGGACATATGAAGTATGCGACAGT 1080
Db 1021 CC-----CTGCACCGAATCGGAAACACACATCAGGATTCTCT 1059
Qy 1081 GGGATGACTTCTCTGATCAGCAAGAGGACACTGACAGCTGTAAGAGTCTGTGAAGTTG 1140
Db 1060 AGGACCCCTGCTGCTGTACAGCGGGGTTTCTTGTGACAAAGATCTCACAATACC 1119
Qy 1141 ATGCTCGCTCAGTGACAGCTTTCGTTCTCCCTCCCATCTTAAATGCCCCAATCTTCAAG 1200
Db 1120 ACAGAGTCTAGACTCGTGTGGACTTCTCTCAATTTTCTAGGGGAGACCCACGTTCC 1179
Qy 1201 AGAGGATGAAGTCTTATAAACTGCACTGATCACTCTTATCTCATTTGTGTGTAGTTTC 1260
Db 1180 TGCCCAAAATTCGAGTCCCAACCTTCAATCACTCACCACCTCTGTCTCTCAATTTG 1239
Qy 1261 TCGTGCCCATCATTTGGCATAGTGGCAGCTCAGCTCTCTGAAATGGGAAACGAAAGTTGCA 1320
Db 1240 TCGTGCTATC-----GCTGGATGTCTCTGCGCGCTTTTATC 1276
Qy 1321 CGGTTGGCTCAGTTAATGACAGATATATCTCCAAGTCCGGAAGGCAAGGAAATGGCAGTG 1380
Db 1277 ATATTCTCTTCACTCTGCTGCTATGCTCTATCTTCTTGTGTTCTTCTGGAATACCA 1336
Qy 1381 AAGATGAAATGAGATTTTCGAGAAGCTGTGATGAAACGATGAGCAACATCGAAAGCAGAA 1440
Db 1337 GGTATGTGCGCGTTTGTCTCTACTTCCAGGAA---CATCACTACAGCAGCGGACCA 1393
Qy 1441 TCAGATATCTTTCAGATAATGAAGCAATCTCTCTAGATGCTAAGAAATTTCCAAATTTCA 1500
Db 1394 TGCAAGACCTTGACGATTCCTGCTCAAGGAACTCTATGTTTCCCTCTT-----GTTGCT 1448
Qy 1501 GCATAACAACTGATCAAGATTTAATGATGTTTCTTTCAGCTAATTTCTTACTTCTTCT 1560
Db 1449 GTACAAAACCTTCGGAACGAACTGCACTTGTATTCCTATCCCATCTCTCGGCTTTTCG 1508
Qy 1561 CCATCCAGGAACATGAGATATCATAGGGGATATCTCCAAGTCAATTAGTAGTCTGAACA 1620
Db 1509 CAAGATTCCTATGGGAGTGGGCTCAGTCCGTTTCTCTCTG----- 1549
Qy 1621 CCAAGATCTTGTATTTGCGAGTTTCAGTATTGAAACACTGAAATGGCAGAGTCCAGAGAAATG 1680
Db 1550 -CTCAGTTTACTAGTGCCATTTGTTGTTCAAGTGGTTC---GTAGGGCTTTTCCCGGCTGTTG 1605

Db 3060 TCCGGCTGTCAGCGCAGGGGCGCCGGTCTCTTTTGTCAAGACCGACCTGTCGGGTGCC 3119
Qy 3896 TGAATGAATCTCAGAGCAGGCGCGCTATCGTGGCTGCGCCACGACGGCGCTTCTTT 3955
Db 3120 TGAATGAATCTCAGAGCAGGCGCGCTATCGTGGCTGCGCCACGACGGCGCTTCTTT 3179
Qy 3956 CGCAGCTGTCTCGACGCTTCTCACTGAAGCGGGAAGGGAATGCGCTGCTATTTGGCGGAAG 4015
Db 3180 GCGCAGCTGTCTCGACGCTTCTCACTGAAGCGGGAAGGGAATGCGCTGCTATTTGGCGGAAG 3239
Qy 4016 TCGCGGGCAGGATCTCTGTCACTCACTTGTCTCGCGAGAAAGTATCCATCATGG 4075
Db 3240 TCGCGGGCAGGATCTCTGTCACTCACTTGTCTCGCGAGAAAGTATCCATCATGG 3299
Qy 4076 CTGATGCATCGCGGGCTGATACGCTTGTATCGGCTACCTGCCCATTCGACCAACCAAG 4135
Db 3300 CTGATGCATCGCGGGCTGATACGCTTGTATCGGCTACCTGCCCATTCGACCAACCAAG 3959
Qy 4136 CGAAACATCGCATCGAGCGACAGTACTCTCGATGGAAGCGGCTTGTGATCAGGATG 4195
Db 3360 CGAAACATCGCATCGAGCGACAGTACTCTCGATGGAAGCGGCTTGTGATCAGGATG 3419
Qy 4196 ATCTGGACGAAGAGCATCAGGGCTCGCGCAGCGCAACTGTTCGCCAGGCTCAAGCGC 4255
Db 3420 ATCTGGACGAAGAGCATCAGGGCTCGCGCAGCGCAACTGTTCGCCAGGCTCAAGCGC 3479
Qy 4256 GCATGCCCGACGGCGAGGATCTCGTGTGACCCATGCGGATGCGGCTTGTGCGGAATATCA 4315
Db 3480 GCATGCCCGACGGCGAGGATCTCGTGTGACCCATGCGGATGCGGCTTGTGCGGAATATCA 3539
Qy 4316 TGGTGGAAATGGCGGCTTTCTGGAATCATCGACTGTGGCGGCTGGGTGTGGCGAC 4375
Db 3540 TGGTGGAAATGGCGGCTTTCTGGAATCATCGACTGTGGCGGCTGGGTGTGGCGAC 3599
Qy 4376 GCTATCAGGACATAGCGTGTGCTACCGGTGATATGCTGAAGAGCTTGGCGGGAATGGG 4435
Db 3600 GCTATCAGGACATAGCGTGTGCTACCGGTGATATGCTGAAGAGCTTGGCGGGAATGGG 3659
Qy 4436 CTGACCGGCTCTCGTGTCTTACGGTATCGGCTATCGGCTTCCGATTCGACGGCATCGCTTCT 4495
Db 3660 CTGACCGGCTCTCGTGTCTTACGGTATCGGCTTCCGATTCGACGGCATCGCTTCT 3719
Qy 4496 ATCGGCTCTTGAAGAGTCTTCTGAGCGGACTCTGGGGTTCGATA 4542
Db 3720 ATCGGCTCTTGAAGAGTCTTCTGAGCGGACTCTGGGGTTCGATA 3766

RESULT 12
AX382149 5130 bp DNA linear PAT 18-MAR-2002
LOCUS
DEFINITION Sequence 9 from Patent WO0202738.
ACCESSION AX382149
VERSION AX382149.1 GI:19576958
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Bremel,R.D., Miller,L.U., Bleck,G.T. and York,D.
TITLE Host cells containing multiple integrating vectors
JOURNAL Patent: WO 0202738-A 9 10-JAN-2002;
Gala Design, Inc. (US)
FEATURES
source
1. 5130
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"

ORIGIN

Query Match 30.7%; Score 1588.6; DB 6; Length 5130;
Best Local Similarity 66.1%; Pred. No. 0;
Matches 3007; Conservative 0; Mismatches 754; Indels 786; Gaps 18;

Qy 1 TTTGAAGACCCCAACCGTAGGTGGCAAGCTAGCTTAAGTAAACGCCACTTTTGCAGGCAT 60
Db 1 TTTGAAGACCCCAACCGTAGGTGGCAAGCTAGCTTAAGTAAACGCCACTTTTGCAGGCAT 60
Qy 61 GGAATAATACATTAATCTAGGAATAGAAAGTTTCAGATCAAGGTCAAGGCAAGAAACAGC 120
Db 61 GGAATAATACATTAATCTAGGAATAGAAAGTTTCAGATCAAGGTCAAGGCAAGAAACAGC 120
Qy 121 TGAATACCAACACAGGATATCTGTGTAAGCGGTTCCTGCCCGGCTCAGGGCCCAAGACA 180
Db 121 TGAATACCAACACAGGATATCTGTGTAAGCGGTTCCTGCCCGGCTCAGGGCCCAAGACA 180
Qy 181 GATGAGACACGCTGATGATGGGCCAAACAGGATATCTGTGTAGCAGTTCCTGCCCGG 240
Db 181 GATGAGACACGCTGATGATGGGCCAAACAGGATATCTGTGTAGCAGTTCCTGCCCGG 240
Qy 241 CTCGGGSCCAAGAACAGATGGTCCCAGATGCCGTCCAGCCCTCAGCAGTTTCTAGTGA 300
Db 241 CTCGGGSCCAAGAACAGATGGTCCCAGATGCCGTCCAGCCCTCAGCAGTTTCTAGTGA 300
Qy 301 TCATCAGATGTTTCCAGGGTCCCCAAGGACCTGAAAATGACCTGTACCTTATTTGAAC 360
Db 301 TCATCAGATGTTTCCAGGGTCCCCAAGGACCTGAAAATGACCTGTACCTTATTTGAAC 360
Qy 361 TAACCAATCAGTTTCGGCTTCTCGCTTCTGTTCGGCGCTTCCGCTCTCCGAGTCAATAA 420
Db 361 TAACCAATCAGTTTCGGCTTCTCGCTTCTGTTCGGCGCTTCCGCTCTCCGAGTCAATAA 420
Qy 421 AGAGCCCAACCCCTCACTCGCGCGCCAGTCTTCGATAGACTGCGTCCCGCGGTAC 480
Db 421 AGAGCCCAACCCCTCACTCGCGCGCCAGTCTTCGATAGACTGCGTCCCGCGGTAC 480
Qy 481 CCGTATTCGCAATAAGCCCTTCTGCTTTCGATCCGATCGTGGTCTCCGCTTCTCTTG 540
Db 481 CCGTATTCGCAATAAGCCCTTCTGCTTTCGATCCGATCGTGGTCTCCGCTTCTCTTG 540
Qy 541 GGAAGGTCTCTCTGAGTGAATGACTACCAACGAGGGGTCTTTTCAATTTGGGGGTCTG 600
Db 541 GGAAGGTCTCTCTGAGTGAATGACTACCAACGAGGGGTCTTTTCAATTTGGGGGTCTG 600
Qy 601 CCGGGAATTTGGAGACCCCTGCCAGGAGCACACCCACCCAGGAGTAAAGTGGCC 660
Db 601 CCGGGAATTTGGAGACCCCTGCCAGGAGCACACCCACCCAGGAGTAAAGTGGCC 660
Qy 661 AGCAACTTATCTGTGCTGCTGATGCTAGTGTCTATGTTTGAATGCGCTTCTGCG 720
Db 661 AGCAACTTATCTGTGCTGCTGATGCTAGTGTCTATGTTTGAATGCGCTTCTGCG 720
Qy 721 TCTGTACTAGTAACTAGCTAGCTCTGATCTCGCGGACCCGCTGGGAACTGACGAGTT 780
Db 721 TCTGTACTAGTAACTAGCTAGCTCTGATCTGATCTCGCGGACCCGCTGGGAACTGACGAGTT 780
Qy 781 CTGAACACCCCGCGCAACCCCTCGGAGACGCTCCAGGGAATTTGGGGCCGTTTCTGG 840
Db 781 CTGAACACCCCGCGCAACCCCTCGGAGACGCTCCAGGGAATTTGGGGCCGTTTCTGG 840
Qy 841 CCGGACTTGAAGAGGAGTGCATGCGAATCCGACCCCGCTCAGGATATGTTGTTCTGGT 900
Db 841 CCGGACTTGAAGAGGAGTGCATGCGAATCCGACCCCGCTCAGGATATGTTGTTCTGGT 900
Qy 901 AGCAGACGAGAACTTAAACAGTTCCGCTCTCGTGAATTTTGTCTTTCGGTTTCGAA 960
Db 901 AGCAGACGAGAACTTAAACAGTTCCGCTCTCGTGAATTTTGTCTTTCGGTTTCGAA 960
Qy 961 CCGAAGCCGCGCTTCTGCTGCGAGCAAGCTTGGGCTGCGAGTTCGACTCTAGAGGA 1020
Db 961 CCGAAGCCGCGCTTCTGCTGCGAGCAAGCTTGGGCTGCGAGTTCGAGGAGCTGGGA 1020
Qy 1021 TCAATTCGGCACGAGTAAATCGGTGCTCCGCTCTTTAGGACATATGAAGTATGGCAGT 1080
Db 1021 CC-----CTGCACCGAATCGGAGAACACATCAGGATTCCT 1059

| | | | |
|----------------------------|------|---|------|
| misc_recomb | | /protein id="AAA72065.1" | |
| LTR | | /db_xref="GI:208849" | |
| misc_recomb | | /translation="MIEDGLHAGSPAAWVERLFGYDWAQOTICSDAAVFLSAQGR PVLVKTDLGALNELQDEARLRLWLTATGVPCAAVLDVVTVEARGLLLEVPQDL LSSHLAPAEKVSIMADAMRRLHTLPATCPDHOAKHRIERARTMEAGLVQDDLOE HQGLAPAEFLARKMPDCEDELVTHGDACLINIMVENGREPSGFI DCGRLGVADRY BDIALATEDIAEELGGEWADRPVLVYGLIAPDSQRIAFYRLLLDEFF" 2877*2878 | |
| | | /note="Tn5 DNA end-Mo-MuLV DNA start" | |
| | | 2920..3513 | |
| | | /note="Mo-MuLV 3' long terminal repeat" | |
| | | 3582*3583 | |
| | | /note="Mo-MuLV DNA end-plasmid pBR322 DNA start" | |
| ORIGIN | | | |
| Query Match | | 28.7%; Score 1486; DB 11; Length 5874; | |
| Best Local Similarity | | 100.0%; Pred. No. 0; | |
| Matches 1486; Conservative | | 0; Mismatches 0; Indels 0; Gaps 0; | |
| Qy | 3692 | GATCTGATCAAGAGACAGGATGAGATCGTTTCGCATGATTGAACAAGATGGATGCAAG | 3751 |
| Db | 2031 | GATCTGATCAAGAGACAGGATGAGATCGTTTCGCATGATTGAACAAGATGGATGCAAG | 2090 |
| Qy | 3752 | CAGTTCTCCGGCCGCTTGGGTGGAGAGGCTATTTCGGCTATGACTGGGCACACAGACAA | 3811 |
| Db | 2091 | CAGTTCTCCGGCCGCTTGGGTGGAGAGGCTATTTCGGCTATGACTGGGCACACAGACAA | 2150 |
| Qy | 3812 | TCGGCTGCTCTGATGCGCCGCTGTTCCGGCTGTGAGCGCAGGGCGCCGGTCTCTTTTG | 3871 |
| Db | 2151 | TCGGCTGCTCTGATGCGCCGCTGTTCCGGCTGTGAGCGCAGGGCGCCGGTCTCTTTTG | 2210 |
| Qy | 3872 | TCAAGACCGACCTGTCCGGTCCCTCGAATGAATGCAGGACGAGCGCGGCTATCGT | 3931 |
| Db | 2211 | TCAAGACCGACCTGTCCGGTCCCTCGAATGAATGCAGGACGAGCGCGGCTATCGT | 2270 |
| Qy | 3932 | GGCTGGCCACACCGGGCTTCTTGGCGAGCTGTCTCGAGCTGTGTCATGAAAGCGGGAA | 3991 |
| Db | 2271 | GGCTGGCCACACCGGGCTTCTTGGCGAGTGTCCGGCGAGGATCTCTCTCATCTCACCTTGC | 2330 |
| Qy | 3992 | GGGACTGGCTCTTATGGGCGAAGTGGCGGGCAGGATCTCTGTCTCATCTCACCTTGC | 4051 |
| Db | 2331 | GGGACTGGCTCTTATGGGCGAAGTGGCGGGCAGGATCTCTCTCATCTCACCTTGC | 2390 |
| Qy | 4052 | CTGCCGAGAAAGTATCCATCATATGCTGATGCAATGCGCGGCTGCATACGCTTGATCCGG | 4111 |
| Db | 2391 | CTGCCGAGAAAGTATCCATCATATGCTGATGCAATGCGCGGCTGCATACGCTTGATCCGG | 2450 |
| Qy | 4112 | CTACTGCGCATTTGACACCAACAGGAAACATCGCATGCGAGCGAGCTACTCGATGG | 4171 |
| Db | 2451 | CTACTGCGCATTTGACACCAACAGGAAACATCGCATGCGAGCGAGCTACTCGATGG | 2510 |
| Qy | 4172 | AAGCGGTCTTGTGATCAGATGATCTGGACGAGAGCATCAGGGGCTCGCGCCAGCCG | 4231 |
| Db | 2511 | AAGCGGTCTTGTGATCAGATGATCTGGACGAGAGCATCAGGGGCTCGCGCCAGCCG | 2570 |
| Qy | 4232 | AAGTGTTCGCGAGGCTCAAGCGGGCAGTCGCCGAGGATCTCGTCTGACCCCATG | 4291 |
| Db | 2571 | AAGTGTTCGCGAGGCTCAAGCGGGCAGTCGCCGAGGATCTCGTCTGACCCCATG | 2630 |
| Qy | 4292 | GCGATGCTTTCGCGAATATCATGFTGGAAATGGCCGCTTTTCTGGATTCATCGACT | 4351 |
| Db | 2631 | GCGATGCTTTCGCGAATATCATGFTGGAAATGGCCGCTTTTCTGGATTCATCGACT | 2690 |
| Qy | 4352 | GTGGCGGCTGGGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCGGTATATG | 4411 |
| Db | 2691 | GTGGCGGCTGGGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCGGTATATG | 2750 |
| Qy | 4412 | CTGAAGAGCTTGGCGGGAATGGGCTGACCGCTTCTCTGCTTTTACGGTATCGCCGCTC | 4471 |
| Db | 2751 | CTGAAGAGCTTGGCGGGAATGGGCTGACCGCTTCTCTGCTTTTACGGTATCGCCGCTC | 2810 |
| Qy | 4472 | CCGATTGCGACGCGATCGCCCTTCTATCGCCCTTTCTGACGAGTTCTTCTGACGCGGACTCT | 4531 |

Search completed: February 27, 2006, 16:11:36
Job time : 25052 secs

THIS PAGE BLANK (USPTO)